Merging Quadratic Programming with Kernel Smoothing for Automated Cluster Expansions of Complex Lattice Hamiltonians

by

Osman Burak Okan

B.S. Materials Science and Engineering, Sabanci University (2004)

M.Sc. Materials Science and Engineering, Sabanci University (2005)

Submitted to the Department of Materials Science and Engineering in partial fulfillment of the requirements for the degree of Master of Science in Materials Science and Engineering at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2008

Author . . . . . .

Department of Materials Science and Engineering

September 21, 2007

Certified by. . . . . Oct 24, 2007

Gerbrand Ceder, R. P. Simmons Professor of Materials Science,

Thesis Supervisor

Accepted by . . . . . .

Samuel M. Allen, POSCO Professor of Physical Metallurgy Chair,

Department Committee on Graduate Students

Merging Quadratic Programming with Kernel Smoothing for Automated Cluster Expansions of Complex Lattice Hamiltonians

by

Osman Burak Okan

Submitted to the Department of Materials Science and Engineering in partial fulfillment of the requirements for the degree of Master of Science in Materials Science and Engineering

Abstract

We present a general outline for automating cluster expansions of configurational energetics in systems with crystallographic order and well defined space group symmetry. The method presented herein combines constrained optimization techniques of positive-definitive quadratic forms with the mathematical tool of Tikhonov regularization (kernel smoothing) for automated expansions of an arbitrary general physical property without compromising the underlying physics. Throughout the thesis we treat formation energy as the fundamental physical observable to expand on since the predominant application of cluster expansions is the extraction of robust approximations for configurational energetics in alloys and oxides. We therefore present the implementational aspects of the novel algorithmic route on a challenging material system Na_xCoO_2 and reconstruct the corresponding GGA ground state line with arbitrary precision in the formation energy-configuration space. The mathematical arguments and proofs, although discussed for cases with arbitrary spin assignments and multiple candidate species for single site occupancy, are eventually formulated and illustrated for binary systems.

Various numerical challanges and the way they are resolved in the framework of kernel smoothing are addressed in detail as well. However, the applicability of the procedure described herein is more universal and can be tailored to probe different observables without resorting to modifications in the algorithmic implementation or the fundemantal mathematical construction. The effectiveness in recovering correct physics shall than be solely tied to the presence of superposable nature (of the physical property of interest) of local atomic configurations or lackthereof.

Acknowledgements

I am nothing short of being extremely indebted to Prof. Ceder for his infinite patience in accomodating my stubborn insistence in pursuing fancy math rather than dissipating my energy in deciphering the physics of the material system at hand.

Now that my personal explorations culminated in a single working algorithm, to best of my hope, the end result might at least partially help me in paying this personal debt to him. After a relentless fight of about two years against a hefty opponent, I can state with absolute confidence that his constant single remark from the first day, indeed turned out to be the ultimate strategy for carrying out reliable cluster expansions; "you have to look at the ground states!"

The other two individuals deserving much of my appreciation are Timothy Mueller and Chris Fischer from whom I learned pretty much all the aspects of cluster expansions and the tricks of the trade. Their comments and insights were the quintessential intellectual ingredients in building up of my own perspective for this physico-mathematical tool which is mastered by only a small clique of scientists. I am also indebted to each and every reamining person in our lab for a plethora of reasons but in particular for making up such a nice environment to do science in.

My father, Ahmet Okan, has always been a very special influence in my life. A large chunk of my scientific noesis directly follows from his expectionally cogent answers to my unending questions. Needless to say my obsession with rigorous mathematical reasoning is a direct result of my childhood spent in his close proximity. I could not have been any luckier dad!

Finally all my work so far and this thesis in particular is dedicated to my fiancee Oznur Tastan, who has never ceased to put her faith in me; even on multiple occasions in which I eventually did...

Contents

List of Tables

List of Figures

Chapter 1

Introduction and Motivation

In materials science, we are fundamentally concerned with the structure, its implications on pertinent physical properties and the resulting macroscopic phenomenology. The most general picture regarding a particular material system, in this respect comes from its phase diagram and physical variables/observables governing the phase stability. Atomistic modelling tools, in this regard, provide invaluable microscopic insight while broadening the extent of experimental means noticeably. As the two most fundamental entities tractable in computational modelling are the configuration and dynamics of atomic species, laying the microscopic principles of atomic ordering as well as the identification of governing thermodynamical principles conveniently falls within the reach of a theorist.

From a theoretical point of view, the biggest challenge is to be able to resolve the underlying physical principles with desired accuracy and exploit them in guiding our configurational search (via spin-flip Monte Carlo simulations) with a fitted Hamiltonian of a particular functional form (polynomial, harmonic etc.). In this work we lay the foundation for an automated expansion of any function of configuration on a complex lattice system and illustrate the working principles with the particular working example of an intercalated transition metal oxide, Na_xCoO_2 and its configurational energetics from GGA input.

We start with presenting a detailed account of the cluster expansion procedure and it's fundamental linkage to the ubiquitous Lenz-Ising model. The basic motivations, the fundamental role of symmetry in extending its scope and the pertinent mathematical background is fully laid out before introducing actual algorithms and regression models concerning the numerical aspects of actual implementation. This is followed by a discussion of the methods for accurate ground state line reproduction via the techniquies of constrained optimization via quadratic programming and the imposition of physics through the mathematical tools of regression which are otherwise targeted for broader families of mathematical problems. As the predominant application of cluster expansions is the calculation of first principles phase diagrams; we shall treat the formation energy as the fundamental entity to expand on throughout the thesis and the illustration of the actual implementation will be in accordance with this choice. However, the kernel of the procedure described in the following sections is more universal and can be trivially tailored to probe different observables as long as they possess a locally additive nature of atomic configurations.

Chapter 2

The Lenz-Ising Model

2.1 The Original Model

The original nearest neighbor Lenz-Ising model is based on the interactions among spins decorating a periodic space lattice. The neighboring parallel or antiparallel spins are assigned an effective potential and considered a constant. Combined with a chemical potential term, the total Lenz-Ising energy is written as;

$$E = \underbrace{-\sum_{i}\sum_{j} U\sigma_i\sigma_j}_{\text{sum over Nearest Neighbors}} + \mu H \sum_{i=1}^{N} \sigma_i \tag{2.1.1}$$

The geometry of underlying lattice enters the microscopic description via the coordination number γ, i.e. the number of terms in nearest neighbor summation are given as $\frac{\gamma N}{2}$. This model has actually been proposed by Ising's supervisor Willhelm Lenz in 1920[7, 9, 18]. It was an attempt to bring out a microscopic description of the ferromagnetic phenomena in certain elements like Fe and Ni. Ernst Ising's main contribution was to come up with the exact solution in 1-D. Following this accomplishment his interest has dwindled on spin models after he observed that the solution for a 1-D system with a finite range of interactions would not exhibit a phase transition. Moreover, he construed that adding more dimensions will not alter his conclusion that there would not be a transition. For the others who dared to tackle the simplest problem in 2-D without the presence of an external magnetic field, the routes to solution were still daunting due to the mathematical complexities involved. Finally, after a couple of decades; the exact solution was presented by the noted chemist Lars Onsager in a meeting of the New York Academy of Sciences in 1942. Indeed, the answer turned out to be rather mathematically involved as he had to exploit ideas from hypercomplex analysis (Onsager based his solution on the the tensor product of quaternion algebras) [7, 27]. Albeit its complexity, the exact solution in 2-D sparked another wave of interest in the Lenz-Ising model as a promising theory of the cooperative phenonema. From a purely physical perspective, the non-trivial nature of the lattice Hamiltonian and the resulting phase transformation which can be tracked back onto it, was the primary stimulus of excitement. Today, it is still one of the rare qualified "yes" answers for the very question of whether we can predict the phase transitions given the exact form of a Hamiltonian.

Onsager's solution have been simplified later by Kaufman with the help more elegant and advanced mathematics techniques like spinors and Lie algebras [7]. Today its physical and mathematical foundations are well established and still relevant as it is essentially a starting point for modelling surface phenomena like adsorption.

The 3-D extension of the Lenz-Ising model, as the most relevant variant of the original is also the most challenging, yet the proper study of bulk phenomena entails the accounting of the third dimension. After countless mathematical excursions spanning decades on the 3-D problem behind Onsager's legacy, by year 2000, Sorin Istrail announced that the solution of the general Ising Model in 3-D was NP complete [20]. His method also implied that the main problem was arising from the loss of planarity of the lattice by adding dimensionality. However, there are still open sub-problems like the original Ising Model with equal coupling constants that might turn out to be tractable in polynomial time.

The Ising-like spin decoration for representing binary (can be generalized to ternary systems and continuous spin states) occupation is also the natural framework for the study of cooperative phenomena on real material systems. Nevertheless, the relevant real life problems are far more complicated and subtle than those bearing analytical solutions. For a complex material such as a transition metal oxide; it is not even possible to attain a closed form expression for the hamiltonian by any theoretical/computational means. The absence of analytical routes for complex crystalline materials thus necessiates approximate solutions. Validity of these approaches shall largely depend on the unveiling of the governing physical principles. Luckily with the advent of ab-inito methods, we now have a virtual labratory of materials which provides invaluable quantitative information about the physical observables like the formation energy and heat capacity. With such a tool at hand; the lattice Hamiltonians can now be approximated with enough detail to recover the relevant physics for further predictions. Obviously the most fundamental knowledge of a material system is the phase diagram and materials scientists hone the quantified insights and knowledge from lattice Hamiltonians with the ultimate goal of computing them.

The cluster expansion method provides a computationally tractable and compact way to parameterize the structural energy in the space of atomic configurations using the input energies of selected structures. In principle, any physical property; be it scalar, vectorial or tensorial, can be cluster expanded. The validity of such an expansion will be determined by the additive nature of different local configurational contributions or lack thereof. The local nature of the structural energy and it's decomposibility into the simpler terms of atomic ordering thus renders it as a good candidate to expand on. From a mathematical point of vantage, over a complete inner product space any Hamiltonian can be represented by an infinite series to an arbitrary precision [21, 39]. However, the truncation of expansion with a rather small set of elements is inevitable from a computational point of view as the computational demand for the calculation of all possible decorations on a parent lattice is prohibitively large. Yet, with a rapidly converging series, this poses no problem for the capturing of underlying physics. The real challenge then is to be able to choose an interaction set which will work for the composition range of interest and good at retrieving ground states and the lower energy excitations.

In the following sections we describe a possible generalization of the Lenz-Ising model as the cluster expansion method and establish a mathematical formalism for describing lattice occupancy on periodic lattice systems with complicated interactions.

2.2 Generalization and Mathematical Properties

The generalization of the Ising model entails the construction of an appropriate basis. The natural route to construct a basis to expand any function of configuration is to build up a composite space using the tensor product of point basis. The point basis is represented by an M-tuple of linearly independent functions for M different species that might occupy a particular lattice site. The easiest and intuitive choice is to use the succesive powers of the spin variable upto the power $M-1$. A more complicated choice will also work as long as the linear independence is established with respect to a particular inner product. The construction with succesive powers of the spin variable correspondance is given below;

$$\mathbf{Basis}_{ijk...} = \{1, \phi_1^{ijk...}(\sigma_{ijk...}), \phi_2^{ijk...}(\sigma_{ijk...}), ..., \phi_M^{ijk...}(\sigma_{ijk...})\} = (1, \sigma, \sigma^2, ..., \sigma^{M-1}) \tag{2.2.1}$$

Taking direct product over N lattice points comprising the lattice we end up with a general basis to expand any property as a function of configuration.

$$\mathbf{Basis}_{Lattice} = \mathbf{Basis}_{ijk...} \otimes \mathbf{Basis}_{lmn...} \otimes ... \tag{2.2.2}$$

In this notation the number of running indices enumerating the lattice sites will be equivalent to the dimension of the real lattice one is operating on. For instance points on a 3-D lattice are represented with three indices *ijk*. This notation can be contracted to a single index for enumerating all the points in the lattice with any given order. When the indicial enumeration is compactified this way, the corresponance between the former and latter notation simply becomes; $\mathbf{Basis}_{ijk...} = \mathbf{Basis}_i$. The running single index shall now span a range from 1 to N; the total number of points on the lattice. Using this notation **eqn-2.2.2** can also be rewritten as;

$$\mathbf{Basis}_{Lattice} = \bigotimes_{i=1}^{N} \mathbf{Basis}_i \tag{2.2.3}$$

If there are more than one symmetrically different lattice sites to be considered which may be occupied, the natural way to construct the full lattice basis is by taking the direct product of the sub-lattice point products. This way one can either include the effect of symmetrically different sites (with independent site energies) for a single species' occupancy or a coupling between two sublattices of different occupancy. We can further expand this idea to operate in a composite space in which one of the sublattices is the reciprocal lattice or to couple a discrete basis with a real space sublattice described by continuous spin occupancies. The math involving either procedure shall naturally be of the same nature albeit complications. In the contracted indicial notation a composite basis with different sublattice products can be built as;

$$\mathbf{Basis}_{Lattice} = (\bigotimes_{i=1}^{N} \mathbf{Basis}_i) \otimes (\bigotimes_{j=1}^{P} \mathbf{Basis}_j) \otimes ... \tag{2.2.4}$$

Once the configuration space is built up; normalization and orthogonality can easily be established for the cluster functions with an appropriate inner product. We start by choosing spin variables as 1 and -1 for the occupancy and vacancy of a site;

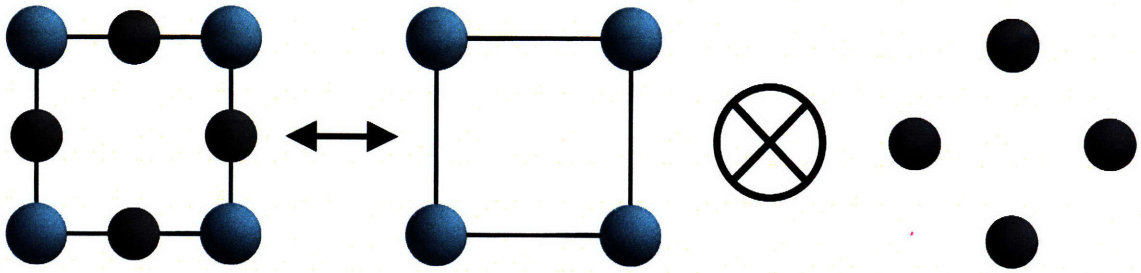

Figure 2.1: An expansion over a composite lattice (left-most) can be decomposed into the tensor product of sub-lattice basis

$$\sigma_{ij} = \begin{cases} 1 & \text{if the lattice site indexed by } ij \text{ is occupied} \\ -1 & \text{if the lattice site indexed by } ij \text{ is vacant;} \end{cases} \tag{2.2.5}$$

Following Sanchez et al. the inner product can be conveniently defined as[29];

$$\langle f(\vec{\sigma}), g(\vec{\sigma}) \rangle = \frac{1}{2^N} \sum_{\vec{\sigma}} f(\vec{\sigma}) g(\vec{\sigma}) \tag{2.2.6}$$

The generic multiplicative term 2^N is the appropriate normalization constant as the inner product sum is taken over all possible configurations decorating a N point lattice. In an alternative notation the summation can be absorbed into a trace sum as;

$$\langle f(\vec{\sigma}), g(\vec{\sigma}) \rangle = \frac{1}{2^N} \mathbf{trace} \left(\sigma_a(\vec{\sigma}) \sigma_b(\vec{\sigma}) \right) \tag{2.2.7}$$

The trace operator, by definition is the sum of diagonal entries for a square matrix. It is possible to justify the usage of a trace in its natural definition by constructing a matrix analogue as well. A general function for a given configuration is shown by $f(\vec{\sigma})$. The fact that each of 2^N configurations can be defined as a vector in a 2^N dimensional space is captured by the conventional vector notation $\vec{\sigma}$. For each configuration there will be a scalar value assigned for $f(\vec{\sigma})$. If we place all those values into vector form the outer product of any such vector pair will yield a $2^N \times 2^N$ matrix.

$$f(\vec{\sigma}_i) \xrightarrow{\text{value of } f \text{ in configuration } i} \begin{bmatrix} f(\vec{\sigma}_1) \\ f(\vec{\sigma}_2) \\ \vdots \\ f(\vec{\sigma}_{2^N}) \end{bmatrix} \tag{2.2.8}$$

$$g(\vec{\sigma}_i) \xrightarrow{\text{value of } g \text{ in configuration } i} \begin{bmatrix} g(\vec{\sigma}_1) \\ g(\vec{\sigma}_2) \\ \vdots \\ g(\vec{\sigma}_{2^N}) \end{bmatrix} \tag{2.2.9}$$

$$\langle f(\vec{\sigma}), g(\vec{\sigma})\rangle = \mathbf{trace}(f(\vec{\sigma}) \otimes g(\vec{\sigma})^T) \tag{2.2.10}$$

$$= \mathbf{trace} \underbrace{\left(\begin{bmatrix} f(\vec{\sigma}_1) \\ f(\vec{\sigma}_2) \\ \vdots \\ f(\vec{\sigma}_{2^N}) \end{bmatrix} \begin{bmatrix} g(\vec{\sigma}_1) & g(\vec{\sigma}_2) & \dots & g(\vec{\sigma}_{2^N}) \end{bmatrix} \right)}_{\mathbf{C}_{(f,g)}} \tag{2.2.11}$$

The trace of this matrix is equivalent our original definition of the inner product between two functions of configuration. In constructing this we actually sum over all possible configurations on a lattice of N atomic sites. Since the concentration is not fixed such a summation is termed as the "Grand Canonical Trace". The vectors $\vec{\sigma}$ assume the role of repeated indices ii in the original definition of trace.

Switching back to the original definition we prove the normalization of cluster functions as;

$$\langle \sigma_a, \sigma_a \rangle = \frac{1}{2^N} \sum_{\vec{\sigma}} \prod_{i \in a} \sigma_i(\vec{\sigma}) \prod_{i \in a} \sigma_i(\vec{\sigma}) \tag{2.2.12}$$

$$= \frac{1}{2^N} \sum_{\vec{\sigma}} \prod_{i \in a} \sigma_i(\vec{\sigma}) \sigma_i(\vec{\sigma}) \tag{2.2.13}$$

$$= \frac{1}{2^N} \sum_{n=1}^{2^N} 1 = 1 \tag{2.2.14}$$

The orthogonality of cluster functions, in a similar manner, is proven as;

$$\langle \sigma_a, \sigma_b \rangle = \frac{1}{2^N} \sum_{\vec{\sigma}} \prod_{i \in a} \sigma_i(\vec{\sigma}) \prod_{j \in b} \sigma_j(\vec{\sigma}) \tag{2.2.15}$$

$$= \frac{1}{2^N} \sum_{\vec{\sigma}} \prod_{i \in a \Delta b} \sigma_i(\vec{\sigma}) \sigma_i(\vec{\sigma}) \prod_{j \in a \cap b} \sigma_j(\vec{\sigma}) \tag{2.2.16}$$

$$= \langle \sigma_{a\Delta b}(\vec{\sigma}), \sigma_{a\Delta b}(\vec{\sigma}) \rangle \sum_{\vec{\sigma}} \prod_{j \in a \cap b} \sigma_j(\vec{\sigma}) \tag{2.2.17}$$

$$= \sum_{\vec{\sigma}} \prod_{j \in a \cap b} \sigma_j(\vec{\sigma}) \tag{2.2.18}$$

$$= 0 \tag{2.2.19}$$

The last step follows from the fact that for a symmetric basis(in general it means we assign a spin variable d for occupancy -d for vacancy) the average concentration for a non-empty cluster will be zero. This can be shown by resorting to the famous Pascal's triangle:

$$
\begin{array}{ccccc}
\binom{0}{0} & & \binom{0}{0} & = 1 = & \langle \sigma(\vec{\sigma})_{\text{ empty}} \rangle \\
\binom{1}{0}\binom{1}{1} & & -\binom{1}{0} + \binom{1}{1} & = 0 = & \langle \sigma(\vec{\sigma})_{\text{ point}} \rangle \\
\binom{2}{0}\binom{2}{1}\binom{2}{2} & \rightarrow & \binom{2}{0} - \binom{2}{1} + \binom{2}{2} & = 0 = & \langle \sigma(\vec{\sigma})_{\text{ pair}} \rangle \\
\binom{3}{0}\binom{3}{1}\binom{3}{2}\binom{3}{3} & & -\binom{3}{0} + \binom{3}{1} - \binom{3}{2} + \binom{3}{3} & = 0 = & \langle \sigma(\vec{\sigma})_{\text{ triplet}} \rangle \\
\vdots & & \vdots & \vdots & \vdots
\end{array}
\tag{2.2.20}
$$

A generic binomial term $\binom{n}{k}$ on the Pascal triangle yields the number of clusters of size n and k sites of which are occupied. Therefore, each binomial term assumes a specific composition. In this setting we observe that the full cancellation of average composition shall only be possible for symmetrical spin assignments. If we choose an asymmetric occupancy set such as $(1, 0)$; the average value of spins will not cancel out as it does for a symmetric set. This implies that, by using asymmetric assignments of spin variables, cluster functions shall only retain the property of linearly independence. The orthogonality is no longer preserved. Therefore a symmetric occupation is a practical choice as well. To illustrate the actual construction; we can use a simple model lattice (with plane group symmetry *p4mm*) comprising only four points. The lattice with its occupation variables enumerated (with two indices)is shown in **fig.2.2**;

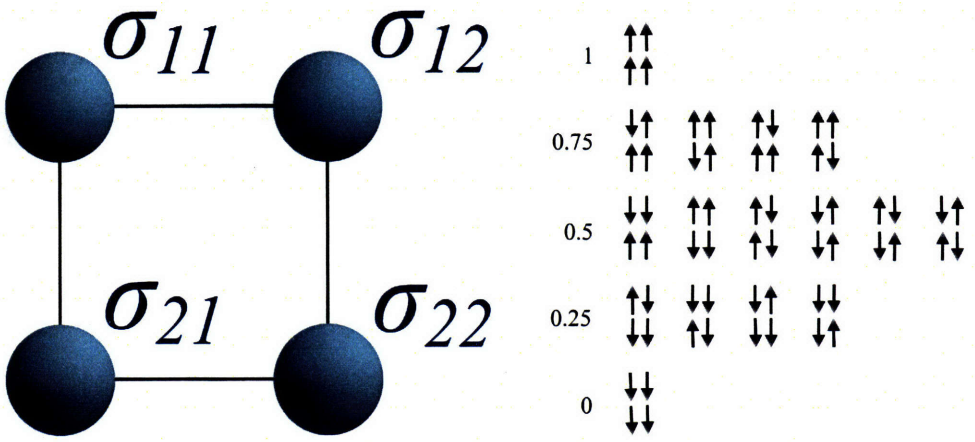

Figure 2.2: Simple model 2-D square lattice for basis construction and the possible decorations of this finite system. The lattice sites are enumerated with a matrix-like convention with two indices: $\sigma = \{\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}\}$ (This construction clearly shows that on a discrete lattice the concentration will also be a discrete function.)

The basis for the product space of four lattice sites is 16 dimensional irrespective of the nature of spin assignments as shown below. It is important to stress the difference between the basis and the spin assignments. The basis is formed in accordance with our choice of how to represent a single lattice point as a function of spin variables. If the functional nature stays same and we merely alter the spin assignments this shall bring out a coordinate transformation in the space of configurations. In the strict sense this is a particular form of a basis change as well. But the correspondence between the space of Chebyshev polynomials and the cluster expansion is established through the functional nature of point assignments

Table 2.1: Group Character Table for the (Nontranslational Part) Plane Group $p4mm$

	E	$A_{\pi/2}$	A_{π}	$A_{3\pi/2}$	σ_{1v}	σ_{2v}	σ_{3v}	σ_{4v}
E	E	$A_{\pi/2}$	A_{π}	$A_{3\pi/2}$	σ_{1v}	σ_{2v}	σ_{3v}	σ_{4v}
$A_{\pi/2}$	$A_{\pi/2}$	A_{π}	$A_{3\pi/2}$	E	σ_{4v}	σ_{1v}	σ_{2v}	σ_{3v}
A_{π}	A_{π}	$A_{3\pi/2}$	E	$A_{\pi/2}$	σ_{3v}	σ_{4v}	σ_{1v}	σ_{2v}
$A_{3\pi/2}$	$A_{3\pi/2}$	E	$A_{\pi/2}$	A_{π}	σ_{2v}	σ_{3v}	σ_{4v}	σ_{1v}
σ_{1v}	σ_{1v}	σ_{2v}	σ_{3v}	σ_{4v}	E	$A_{\pi/2}$	A_{π}	$A_{3\pi/2}$
σ_{2v}	σ_{2v}	σ_{3v}	σ_{4v}	σ_{1v}	$A_{3\pi/2}$	E	$A_{\pi/2}$	A_{π}
σ_{3v}	σ_{3v}	σ_{4v}	σ_{1v}	σ_{2v}	A_{π}	$A_{3\pi/2}$	E	$A_{3\pi/2}$
σ_{4v}	σ_{4v}	σ_{1v}	σ_{2v}	σ_{3v}	$A_{\pi/2}$	A_{π}	$A_{3\pi/2}$	E

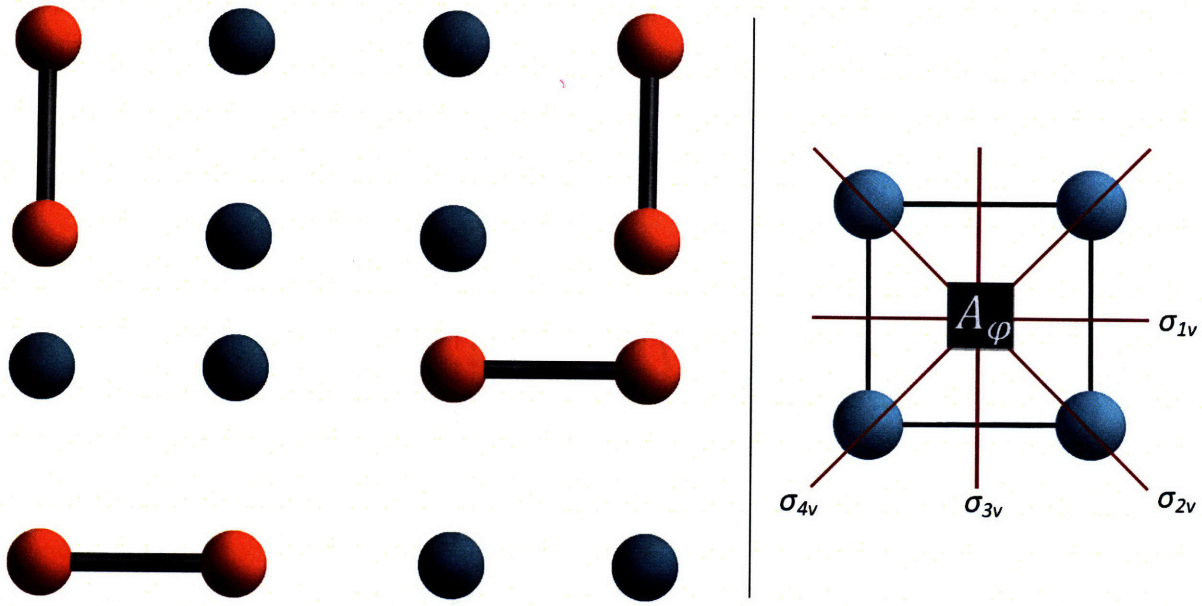

Figure 2.3: The symmetic equivalence and multiplicity between nearest neighbour pairs is shown. It can be established using the symmetry operations in the space group. It is also evident that there are more operations in the space group than the number of symmetrically equivalent clusters; this is a direct result of Lagrange's theorem and the factor group decomposition.

not the particular nature of spin assignments[29]. Therefore, with an arbitrary choice of occupany/vacancy representation the completeness of the expansion is still retained.

$$\begin{bmatrix} 1 \\ \phi_1^{11}(\sigma_{11}) \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \phi_1^{12}(\sigma_{12}) \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \phi_1^{21}(\sigma_{21}) \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \phi_1^{22}(\sigma_{22}) \end{bmatrix} \tag{2.2.21}$$

$$= \begin{bmatrix} 1 \\ \sigma_{11} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \sigma_{12} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \sigma_{21} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \sigma_{22} \end{bmatrix} \tag{2.2.22}$$

$$= \begin{bmatrix} 1 \\ \sigma_{11} \\ \vdots \\ \sigma_{11}\sigma_{12} \\ \vdots \\ \sigma_{11}\sigma_{12}\sigma_{21} \\ \vdots \\ \sigma_{11}\sigma_{12}\sigma_{21}\sigma_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ \sigma_1 \\ \vdots \\ \sigma_1\sigma_2 \\ \vdots \\ \sigma_1\sigma_2\sigma_3 \\ \vdots \\ \sigma_1\sigma_2\sigma_3\sigma_4 \end{bmatrix} \tag{2.2.23}$$

With the knowledge of energies in each of 16 configurations this problem would have reduced to a direct matrix inversion with no fitting error. However the systematic approach of scanning all configurations is plauged by the curse of dimensionality. For this particular example, although direct inversion seems plausible if we have added few more sites, the exact solution would easily fall beyond the reach of computational tractibility. This forces us to be selective in choosing the structures to calculate and correct identification of governing types of interaction.

2.3 Symmetry Reduction and Truncation of Expansion

The spatial symmetry of a crystal is mathematically described via its space group $\mathcal{G}$. In principle, all the symmetrically equivalent positions m can be reached by succesive application of space group elements g to the fractional coordinates of an initial point m_0 [23]. In an ideal crystal extending to infinity, the number of points generated by the group action shall as well be inifinite and they are conveniently collected in an orbit Θ [4, 6, 23]. The elements of this orbit in the crystallographic notation are collectively termed as the Wyckoff family pertatining to the initial position m_0. This is formally represented as;

$$\Theta = \{m \in M \mid m = gm_0 \text{ for some } g \in \mathcal{G}\} \tag{2.3.1}$$

Owing to the spatial symmetry of parent lattices, it is also possible to identify clusters which are equivalent under the action of space group elements. An n-body cluster is a finite subset $A = \{m_1, m_2, ..., m_n\}$ of the underlying lattice M. The symmetric multiplicity is then given by the number of elements in the $\mathcal{G}$ orbit of A.The explicit formulation of such an orbit is anologous to the single point counterpart and formulated in the logical syntax as;

$$\Theta(A) = \{A \subseteq M \mid A = gA_0 \text{ for some } g \in \mathcal{G}\} \tag{2.3.2}$$

Yet another group theoretic concept central to the symmetric reduction process is that of a group normalizer [4]. This notion follows from the basic observation that there shall be sym-

metry operations under which a given cluster will remain globally unaffacted. Consequently, the subset of such elements of the space group is referred to as the group normalizer.

$$\mathcal{N}_A(\mathcal{G}) = \{g \in \mathcal{G} \mid A = gA \text{ for some } A \subseteq M\} \tag{2.3.3}$$

With those terms defined; the multiplicity of a cluster is given by the ratio of the order of the space group(with all the translations are carried back onto the unit cell) to the order of the normalizer. The motivation of tracking such equivalences is the fact that the clusters which are symmetrically the same will assume the same ECI. This brings significant simplification to the expansion and its physical interpretation in terms of configurational contributions. Although we have stated the general algorithm based the concept a mathematical orbit and group normalizer we have to be careful to take into account of possible orbit splitting. This is the reason why there are mutliple fractional positions listed for a certain Wyckoff position and all of them has to be taken into account to ensure proper counting of the symmetric multiplicity of clusters. The symmetry reduction based on space group factorization was first introduced in the context of lattice Hamiltonians by Gratias et al. in an attempt to formalize the configurational entropy. However, they haven't suggest an algorithmic short-cut for this problem.[15] Admittedly, one can propose alternative ways for automated symmetric reduction. The power of this particular route described above lays in its formal description based on concepts from the well established group theoretical nomenclature. The way the orbit splitting is tackled in our work relies on a database of Wyckoff positions to rule out the effect of orbit splitting while adhering to the original route of Gratias et al.

To prove that under a general inhomogeneous transformation the ECIs shall remain invariant, we return back to the inner product definition exploiting the trace operator acting on the matrix $\mathbf{C}_{(f,g)}$. Let us assume that A is an operator which maps two clusters onto each other in the space of fractional coordinates. There will be a higher dimensional analogue of this mapping; $\mathbf{A_{conf}}$ which carries out the corresponding transformation between configuration vectors in the product space. Now the idea is that the trace of matrix $\mathbf{C}_{(f,g)}$ under a similarity transformation imposed by $\mathbf{A_{conf}}$ should remain invariant. This automatically proves the equivalence of ECIs for symmetrically identical clusters;

$$\begin{aligned} \langle f(\vec{\sigma}), g(\vec{\sigma}) \rangle &= \mathbf{trace}(\mathbf{C}_{(f,g)}) && (2.3.4) \\ \mathbf{trace}(\mathbf{A_{conf}}\mathbf{C}_{(f,g)}\mathbf{A}_{\mathbf{conf}}^{-1} &= \mathbf{trace}(\mathbf{C}_{(f,g)}\mathbf{A}_{\mathbf{conf}}^{-1}\mathbf{A_{conf}}) && (2.3.5) \\ &= \mathbf{trace}(\mathbf{C}_{(f,g)}) && (2.3.6) \end{aligned}$$

As we inevitably need to truncate the expansion we seek for rules to follow for adding and subtracting clusters. We cull here the well adopted practices and physical observations for this purpose;

1. Compact clusters are favoured over large clusters as one expects decaying of interaction energy with distance. The way we quantify distance is through the concept of cluster diameter, i.e. the longest distance between a pair of sites in a cluster.

2. Fewer body terms are favoured over many-body terms as they are expected to be lower in energetic magnitude. This is justified by the fact that a coupling between two terms is in general weaker than the individual interactions.

3. A cluster is included if all of it's subclusters are already included. As a corollary to this, a cluster can be removed if there is no supercluster of it ishall be retained in the expansion

The third rule has been proposed for algorithmic purposes and have formerly been discussed elsewhere [37]. The explicit proof of this subcluster necessity, although probably known to many people in the field, has only been recently published[31]. It is essentially a comparison between ECI terms when the spin variables are altered to another set and is completely analogous to the procedure for penalizing unfavourable interactions.

To prove the subcluster rule we first start by noting that the two occupation set can be mapped onto each other by a linear transformation ;

$$\sigma_i = Ax_i + B \tag{2.3.7}$$

Recalling the explicit expression for cluster functions as a product of site occupancies and feeding the transformed values for occupancy we get;

$$\langle \sigma_a(\vec{\sigma}) \rangle = \langle \prod_{i \in a} \sigma_i (\vec{\sigma}) \rangle = \langle \prod_{i \in a} (Ax_i (\vec{\sigma}) + B) \rangle = \sum_{b \subseteq a} A^{n_b} B^{n_a - n_b} \langle \prod_{i \in b}^{n_b} x_i \rangle \tag{2.3.8}$$

Setting up the correspondance between two arbitrary choice of spin occupancies; we rewrite the cluster expanded energy as;

$$\begin{aligned} E &= \sum_a J_a^{\sigma_i} \langle \sigma_a(\vec{\sigma}) \rangle & (2.3.9) \\ &= \sum_a J_a^{\sigma_i} \sum_{b \subseteq a} A^{n_b} B^{n_a - n_b} \langle \prod_{i \in b}^{n_b} x_i \rangle & (2.3.10) \\ &= \sum_b \sum_a J_a^{\sigma_i} A^{n_b} B^{n_a - n_b} \langle \prod_{i \in b}^{n_b} x_i \rangle & (2.3.11) \\ &= \sum_b J_b^{x_i} \langle x_b(\vec{x}) \rangle & (2.3.12) \end{aligned}$$

As an immediate outcome of the last equality; each indidual ECI in one spin assignment for a given cluster comprises all the ECIs of its subclusters (including itself and the empty term with constant correlation)in the other spin assignment. This implies that the subcluster constaint is a mathematical requirement if we want to keep the freedom of switching between arbitrary spin assignments.

The clusters of different sizes in the expansion comprise of site energies(1-body), pair interactions(2-body), and many body terms(n-body). The interpretation of the former two is rather intuitive. A many body term, however, packs more complicated information about the configurational energetics. As an example; a triplet will give information about the interaction of a lattice site and a two-body term(with all the possible occupancies considered). This information is not contained in an expansion which includes only the subclusters of this particular triplet.

The cluster expansion, albeit being a numerical fitting procedure, is inherently governed by the physical characteristics of the material system at hand. The energetics of the lattice

Hamiltonian is coarse grained by neglecting the electronic and vibrational degrees of freedom and the resulting entropy created thereby. With this approximation, the energy differences is solely tied to the configurational changes between distinct lattice decorations and the ECIs will be temperature independent. Once we have an idea of the low energy states this will naturally extend to the dominant terms in the partition function. This way one can rapidly calculate a mean value of a property describing the system using the famous expression from statistical mechanics[3];

$$\langle A \rangle = \mathbf{trace}(\rho A) = \frac{\mathbf{trace}(Ae^{-\beta H})}{\mathbf{trace}(e^{-\beta H})} \tag{2.3.13}$$

Chapter 3

A Novel Composite Route for the Computation of Effective Cluster Interactions

3.1 Computation of the ECI

A generic least squares problem is formulated as the minimization of the Euclidean norm;

$$\arg\min_{x} \|\mathbf{A}\mathbf{x} - \mathbf{b}_{DFT}\|_2 \tag{3.1.1}$$

$\mathbf{A}$ is the mxn correlation matrix for m structures and n clusters including the constant term. From now on we shall name the input vector of DFT energies as $\mathbf{b}_{DFT}$.

We can decompose the non-orthogonal and orthogonal errors of fitting by explicitly rewriting the least squares problem as;

$$\begin{aligned} \|\mathbf{A}(\mathbf{x}+\mathbf{e}) - \mathbf{b}_{DFT}\|_2^2 &= (\mathbf{A}(\mathbf{x}+\mathbf{e}) - \mathbf{b}_{DFT})^T (\mathbf{A}(\mathbf{x}+\mathbf{e}) - \mathbf{b}_{DFT}) && (3.1.2)\\ &= (\mathbf{x}^T\mathbf{A}^T - \mathbf{b}_{DFT}{}^T + \mathbf{A}^T\mathbf{e}^T)(\mathbf{A}\mathbf{x} - \mathbf{b}_{DFT} + \mathbf{A}\mathbf{e}) && (3.1.3)\\ &= \|\mathbf{A}\mathbf{x} - \mathbf{b}_{DFT}\|_2^2 + \|\mathbf{A}\mathbf{e}\|_2^2 && (3.1.4) \end{aligned}$$

The expression above shows that the norm in the least squares problem is minimized if the non-orthogonal component of error(the second term above) is zero[10]. The geometric justification of this method follows from the Pythagoras theorem.

With the orthogonality property of the minimal error is established; we now seek an explicit expression for the minimizer of the problem;

$$\begin{aligned} \|\mathbf{A}\mathbf{x} - \mathbf{b}_{DFT}\|_2^2 &= (\mathbf{A}\mathbf{x} - \mathbf{b}_{DFT})^T (\mathbf{A}\mathbf{x} - \mathbf{b}_{DFT}) && (3.1.5)\\ &= \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} - 2\mathbf{x}^T\mathbf{A}^T\mathbf{b}_{DFT} + \mathbf{b}_{DFT}{}^T\mathbf{b}_{DFT} && (3.1.6) \end{aligned}$$

The first derivative of this expression with respect to $\mathbf{x}$, shall give the global minima since $\mathbf{A}^T\mathbf{A}$ is a positive definitive matrix implying the convexity of the minimized norm[10].

$$\frac{d\|\mathbf{Ax}-\mathbf{b}_{DFT}\|_2^2}{d\mathbf{x}} = \frac{d\left(\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}-2\mathbf{x}^T\mathbf{A}^T\mathbf{b}_{DFT}+{\mathbf{b}_{DFT}}^T\mathbf{b}_{DFT}\right)}{d\mathbf{x}} \tag{3.1.7}$$

$$= 0 \tag{3.1.8}$$

$$= -2\mathbf{A}^T\mathbf{b}_{DFT}+2\mathbf{A}^T\mathbf{Ax} \tag{3.1.9}$$

$$\mathbf{b} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{x} \tag{3.1.10}$$

The second derivative of the minimized norm with respect to $\mathbf{x}$ yields a very simple expression which will be the kernel of the modular method which shall be described shortly;

$$\frac{d^2\|\mathbf{Ax}-\mathbf{b}_{DFT}\|_2^2}{d\mathbf{x}^2} = \frac{d(-2\mathbf{A}^T\mathbf{b}_{DFT}+2\mathbf{A}^T\mathbf{Ax})}{d\mathbf{x}} = \underbrace{\mathbf{A}^T\mathbf{A}}_{\text{HESSIAN}} \tag{3.1.11}$$

However, the importance of matrix $\mathbf{A}^T\mathbf{A}$ is not only its Hessian property.[1] From the least squares analysis, we remember the well-established formula for the variance;

$$\mathbf{VAR}(\hat{\mathbf{b}}_{DFT}) = RMSE(\hat{\mathbf{b}}_{DFT})^2(\mathbf{A}^T\mathbf{A})^{-1} \tag{3.1.12}$$

Yet another closely related property follows from the singular value decomposition(SVD) of the correlation matrix[36, 10];

$$\mathbf{A} = \mathbf{U\Gamma V}^T \tag{3.1.13}$$

Now if we write $\mathbf{A}^T\mathbf{A}$ in it's decomposed form;

$$\mathbf{A}^T\mathbf{A} = \mathbf{V\Gamma U}^T\mathbf{U\Gamma V}^T = \mathbf{V\Gamma}^2\mathbf{V}^T \tag{3.1.14}$$

The last expression is essentially a similarity transformation and the eigenvalue decomposition of $\mathbf{A}^T\mathbf{A}$. Considering that SVD shall always exists (as opposed to the eigendecomposition), we are left out with the fact that $\mathbf{A}^T\mathbf{A}$ shall always have a valid eigendecomposition as it can alternatively be carried out via resorting to the SVD of $\mathbf{A}$. With that said, the SVD of a matrix can be (albeit somewhat loosely due to some technicalities) seen as an eigendecomposition of a related matrix which is once again nothing but $\mathbf{A}^T\mathbf{A}$. This is crucial as we shall see in the description of Tikhonov regularization procedure in the next section. Right before introducing the quadratic programming methods we distill our observations so far into a single statement; in doing a cluster expansion the fundamental entity is the $\mathbf{A}^T\mathbf{A}$ matrix (thus the unifying principle for a modular algorithm).

Now we continue by describing the quadratic programming tool. In accordance with its name it is the general framework involving the optimization of quadratic forms under linear constraints[11, 26].

[1] Hessian $\mathbf{H}$ of a multivariable function is defined as the matrix of second order partial derivatives (or a second rank tensor in more rigorous terms, as matrix notation will not generalize to higher order derivatives). For a quadratic function it is a matrix with constant entries: $\mathbf{H} = \nabla^2 f(x) = \nabla f(x)^T \nabla f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_N \partial x_N} & \cdots & \frac{\partial^2 f(x)}{\partial x_N^2} \end{bmatrix}$

$$\arg\min_{x}(\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x}+\mathbf{c}^T\mathbf{x}) \text{ subject to } \begin{cases} \mathbf{D_{ineq}^T x} \leq \mathbf{f_{ineq}} \text{ with } \mathbf{D}_{ineq}^T = \begin{bmatrix} \vdots \\ -\mathbf{d}_i^T- \\ \vdots \end{bmatrix} ; i \in I \\ \mathbf{D}_{eq}^T\mathbf{x} = \mathbf{f}_{eq} \text{ with } \mathbf{D}_{eq}^T = \begin{bmatrix} \vdots \\ -\mathbf{d}_j^T- \\ \vdots \end{bmatrix} ; j \in \mathcal{E} \end{cases} \tag{3.1.15}$$

Least squares problems inherently belong to this class and their explicit formulation in **eqn-3.1.1** also bears the partial definition of the generic quadratic programming scheme. We first remove the scalar term ${\mathbf{b}_{DFT}}^T\mathbf{b}_{DFT}$ from the formulation as it is only a constant offset value (it only shifts the origin in the problem). We further identify the $\mathbf{Q}$ matrix as $\mathbf{Q} = \mathbf{A}^T\mathbf{A}$, and the $\mathbf{c}^T$ vector as $\mathbf{c}^T = -2\mathbf{A}^T\mathbf{b}_{DFT}^T$.

The biggest difference is the absence of a compact and closed matrix expression as in the constrainted least squares estimator (for the case with inequality constraints). The quadratic program then operates in a constrained domain shaped in accordance wth the way we identify our working constraints and formulate them into a matrix form(i.e the $\mathbf{D}$ matrix). The way those constraints are identified is discussed in the next section.

3.2 Active Set Method for Quadratic Programming

The fundamental entity in a constraint optimization is the Lagrangian of the objective function. The merit of exploiting the concept of Lagrangian is often ignored yet it has one fundamental motivation. The Lagrangian transforms the problem of identifying stationary points of the constrained problem to an equivalent task of unconstrained optimization with additional parameters. Those extra parameters are termed as the Lagrange multipliers and weight the constraints imposed on the objective function. The form of Lagrangian is a statement that not all the constraints affect the objective function around its feasible optima equally strong. For the case of quadratic programming, the Lagrangian is given as;

$$\mathcal{L}(\mathbf{x},\lambda_i) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x}+\mathbf{c}^T\mathbf{x} - \sum_{i\in I\cup\mathcal{E}} \lambda_i \underbrace{(\mathbf{d}_i^T\mathbf{x}-f_i)}_{c_i(\mathbf{x}_i)} \tag{3.2.1}$$

One immediate finding is the need for equality constraints in operating with a Lagrangian therefore we make the distinctive definition of active and inactive constraints. The active set method that will be described in this section fundamentally exploits this distinction between active and inactive set of constraints which shall be operational in reaching the optimal solution. Therefore we precede the algorithmic account of this particular method with a rigorous definition;

$$A(x*) = \{k \in \mathcal{E}\cup I : \mathbf{d}_k^T\mathbf{x} = \mathbf{f}_k\} \tag{3.2.2}$$

The importance of the active constraints lies in the fact that they restrict the dimensionality of search vectors. Such a confined search is compulsory as any move along a feasible direction on the multidimensional landscape has to operate without perturbing the "equality"

conditions. Therefore we the project the objective function gradients and the column space of the Hessian matrix to the complementary space. This algorithm makes full use of the fundamental theorem of linear algebra. [2]

The main challange in problems with inequality constraints stems from the fact that an apriori estimate as to which constraints shall be active at the solution is not readily available. If this can be identified unambigiously, the problem is reduced to a simpler variety of quadratic programming with equality constraints without much effort. At this point we can present the stepwise reduction of the inequality constrained problem as below;

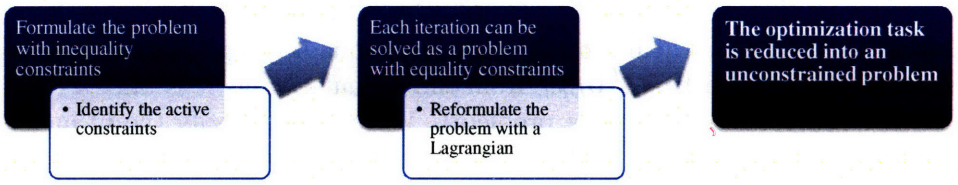

Figure 3.1: Stepwise reduction of the constrained optimization problem

First Order Conditions

First order conditions for constrained optimization closely parallels the concept of a stationary point for the unconstrained case while additionally accounting for the feasibility condition. Feasibility, herein entails that an incremental step in a search direction shall leave the active constraints intact; $c_i(\mathbf{x}^* + \boldsymbol{\delta}) = c_i(\mathbf{x}^*)$ Taylor expanding the constraint function;

$$c_i(\mathbf{x}^* + \boldsymbol{\delta}) = c_i(\mathbf{x}^*) + \boldsymbol{\delta}^T \nabla c_i(\mathbf{x}^*) + o(|\boldsymbol{\delta}|) \tag{3.2.3}$$

we establish the first condition that;

$$\mathbf{s}^T \nabla c_i(\mathbf{x}^*) = \mathbf{s}^T \mathbf{d}_i^* \tag{3.2.4}$$

Therefore we can state that a feasible search direction $\mathbf{s}$ shall be orthogonal to the constraint gradients. If such a point also yields a negative slope in the objective function (again resorting to the Taylor expansion);

$$f(\mathbf{x}^* + \boldsymbol{\delta}) = f(\mathbf{x}^*) + \boldsymbol{\delta}^T \nabla f(\mathbf{x}^*) + o(|\mathbf{s}|) \tag{3.2.5}$$

$$f(\mathbf{x}^* + \mathbf{s}) - f(\mathbf{x}^*) = \mathbf{s}^T \nabla f(\mathbf{x}^*) < 0 \tag{3.2.6}$$

we would improve on the minimization task by moving away from our initial point. Thus we observe that for a local minima the two equalities $\mathbf{eqn} - \mathbf{3.2.4}$ and $\mathbf{eqn} - \mathbf{3.2.6}$ can not be satisfied concurrently. It is obvious that this argument will hold for any objective gradient vector which falls within the span of constraint gradient vectors;

$$\nabla f(\mathbf{x}^*) \in \mathbf{span}(\nabla c_i(\mathbf{x}^*)) \tag{3.2.7}$$

[2] Rank Nullity theorem states that: $dim(Null(\mathbf{A})) + dim(Range(\mathbf{A})) = N$ and $Null(\mathbf{A}) \oplus Range(\mathbf{A}^T) = \mathbb{R}^N$

Once we project the objective function into this subspace, the coefficients shall be nothing other than the corresponding lagrange multipliers;

$$\nabla f(\mathbf{x}^*) = \sum_{i \in E} \lambda_i \nabla c_i(\mathbf{x}^*) \tag{3.2.8}$$

For the case of linear constraints, we have already identified the constraint gradients as the the rows of the constraint matrix $\mathbf{D}$ therefore we can cast this equality into a matrix equation and derive an explicit formula for the Lagrange multipliers;

$$\nabla f(\mathbf{x}^*) = \sum_{i \in E} \lambda_i \mathbf{d_i} = \mathbf{D}\boldsymbol{\lambda} \rightarrow \boldsymbol{\lambda} = (\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}\nabla f(\mathbf{x}^*) \tag{3.2.9}$$

The motivation for Lagrange multipliers is laid out by assessing the effect of perturbations on the Lagrangian through constraints;

$$c_i(x) = 0 \rightarrow c_i(\mathbf{x}) = \boldsymbol{\epsilon} \text{ and } \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \rightarrow \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon}) \tag{3.2.10}$$

Writing the first order differential of the Lagrangian with the perturbations as new function parameters;

$$d\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon}) = d\mathbf{x}^T \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon}) + d\boldsymbol{\lambda}^T \nabla_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon}) + d\boldsymbol{\epsilon}^T \nabla_{\boldsymbol{\epsilon}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon}) \tag{3.2.11}$$

$$\frac{d\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon})}{d\boldsymbol{\epsilon}} = \frac{\partial \mathbf{x}}{\partial \boldsymbol{\epsilon}}^T \underbrace{\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon})}_{\mathbf{0}} + \frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{\epsilon}}^T \underbrace{\nabla_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon})}_{\mathbf{0}} + \nabla_{\boldsymbol{\epsilon}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon}) \tag{3.2.12}$$

$$\frac{d\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon})}{d\boldsymbol{\epsilon}} = \underbrace{\nabla_{\boldsymbol{\epsilon}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon})}_{\boldsymbol{\lambda}} \tag{3.2.13}$$

$$\frac{df(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\epsilon})}{d\boldsymbol{\epsilon}} = \boldsymbol{\lambda} \tag{3.2.14}$$

The concept of Lagrange mutliplier thus works as a quantifying aid for the sensitivity of the objective function to the presence of a particular constraint.

The observations on first order necessary optimality conditions can be collected into a single statement which we call the Kuhn-Tucker conditions[5, 11].

$$\text{Kuhn-Tucker conditions} \begin{cases} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda) = 0 \\ c_i = 0 \text{ for } i \in \mathcal{E} \\ c_i \geq 0 \text{ for } i \in I \\ \lambda_i \geq 0 \text{ for } i \in I \\ \lambda_i c_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{E} \cup I \end{cases} \tag{3.2.15}$$

Any point that satisfies those is termed as a Kuhn-Tucker point. However, the necessary conditions are not only of purely mathematical interest and have a rather strong geometric motivation in linking the objective function gradient with constraint gradients and the

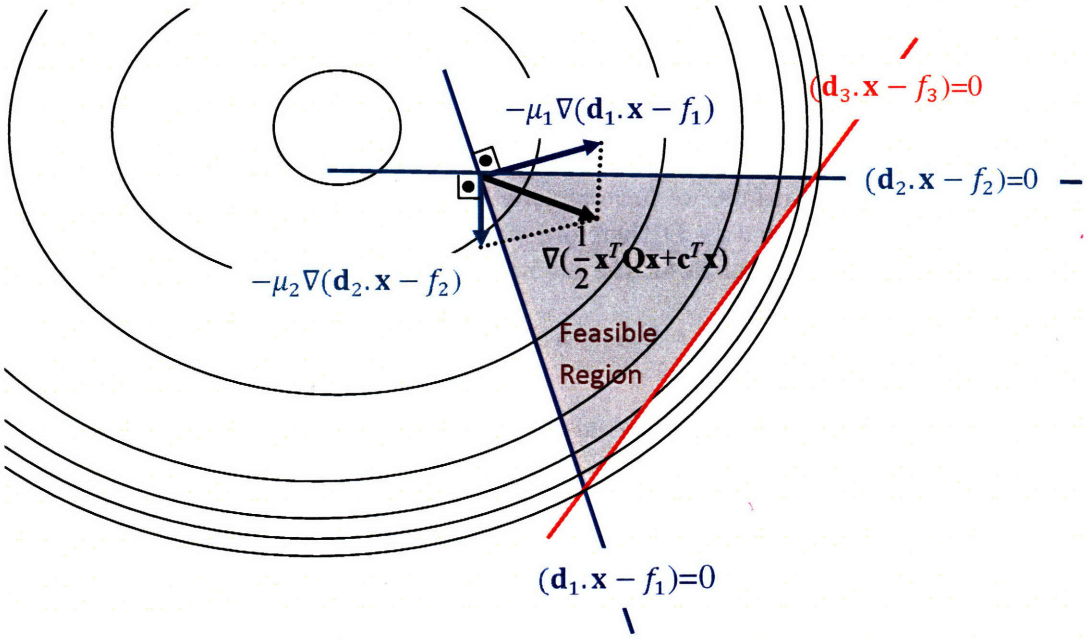

Figure 3.2: Kuhn-Tucker conditions with linear constraints.The general projective coefficients μ_1 and μ_2 are nothing but the Lagrange multipliers for the initial set of active constraints operational at the starting point(where the tails of three arrows meet).

Lagrange multiplers.

Second Order Conditions

Second order conditions provide the vital ingredient of curvature characteristics. The fundamental statement regarding this type of condition is again construed by looking at the Taylor expansion now including upto to the second order terms;

$$f(\mathbf{x}^* + \boldsymbol{\delta}) = \mathcal{L}(\mathbf{x}^* + \boldsymbol{\delta}) \tag{3.2.16}$$

$$= \mathcal{L}(\mathbf{x}^*) + \boldsymbol{\delta}^T\nabla\mathcal{L}(\mathbf{x}^*) + \frac{1}{2}\boldsymbol{\delta}^T\nabla^2\mathcal{L}(\mathbf{x}^*)\boldsymbol{\delta} + O(|\boldsymbol{\delta}|) \tag{3.2.17}$$

$$= f(\mathbf{x}^*) + \frac{1}{2}\boldsymbol{\delta}^T\nabla^2\mathcal{L}(\mathbf{x}^*)\boldsymbol{\delta} \tag{3.2.18}$$

Minimality entails that the value of objective function shall increase when we move from the minima by $\boldsymbol{\delta}$. So for any feasible direction $s$we can establish that;

$$f(\mathbf{x}^* + \mathbf{s}) - f(\mathbf{x}) \geq 0 \rightarrow \mathbf{s}^T\nabla^2\mathcal{L}(\mathbf{x}^*)\mathbf{s} \geq 0 \tag{3.2.19}$$

For a Hessian (which is essentially a quadratic form) the positive-semidefitinitiveness implies a global solution while the positive-definitiveness brings out the additional property of uniqueness[3]. Therefore, on a convex landscape, any local minima will automatically establish

[3]The Hessian pertaining to the Lagrangian of a generic objective function shall be; $\nabla^2\mathcal{L}(\mathbf{x}^*) = \nabla^2 f(\mathbf{x}^*) - \sum_{i\in E}\lambda_i^*\nabla^2 c_i(\mathbf{x}^*)$

itself as the global minima. The positive-definitiveness is inherently linked to the eigenvalue spectrum of the Hessian and is achieved only through a strictly positive spectrum. The other remaining possibility for the form of Hessian is the case for which we would have an indefinite quadratic form which is characterized by the existence of saddle points due to the mixed sign of eigenvalues. Fortunately, the least squares problems are inherently convex since the Hessian is symmetric and positive definitive which immensely alleviates the solution procedure as we take full advantage of a convex topography[4].

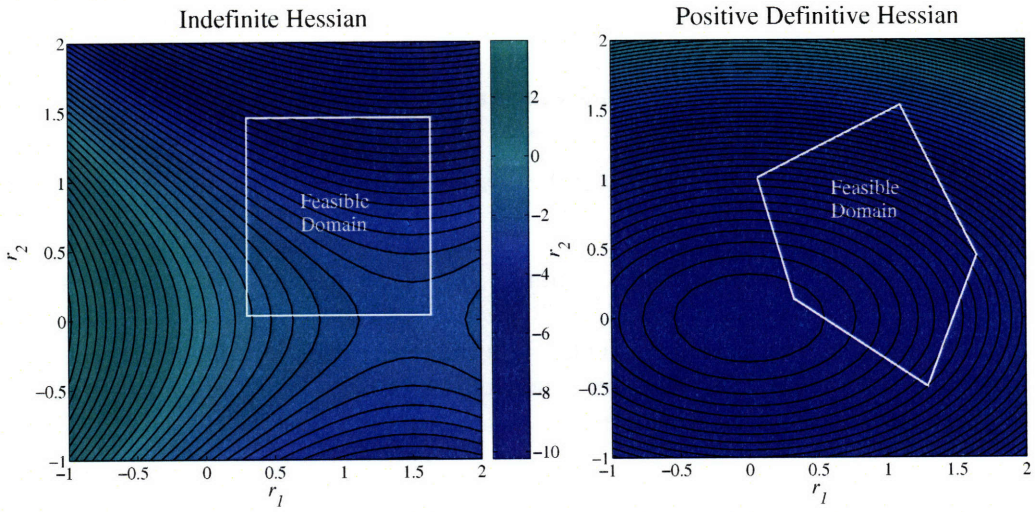

Figure 3.3: The constrained domain of an indefinite problem and a positive-definite (convex) problem

Choosing the feasible step size, along with a feasible direction is the most fundamental step in iterating a quadratic programming algorithm. We can naturally write down the consequitive step updates for the case in which all the constraints in our working set is satisfied as;

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k \tag{3.2.20}$$

The updated solution than would be found from;

$$f(\mathbf{x}_k + \mathbf{p}_k) = (\mathbf{x}_k + \mathbf{p}_k)^T \mathbf{Q}(\mathbf{x}_k + \mathbf{p}_k) + \mathbf{b}^T(\mathbf{x}_k + \mathbf{p}_k) \tag{3.2.21}$$

$$= \underbrace{\frac{1}{2}\mathbf{p}_k^T \mathbf{Q}\mathbf{x}_k + \mathbf{b}^T \mathbf{x}_k}_{\text{Constant term, } r_k} + \frac{1}{2}\mathbf{p}_k^T \mathbf{Q}\mathbf{p}_k + \mathbf{b}^T \mathbf{p}_k \tag{3.2.22}$$

$$= \frac{1}{2}\mathbf{p}_k^T \mathbf{Q}\mathbf{p}_k + \mathbf{b}^T \mathbf{p}_k + r_k \tag{3.2.23}$$

Therefore we can readily identify the feasible direction pertaining to the k^{th} iterate by solving the pertinent Kuhn-Tucker system;

[4]This can also be understood by the fact that the covariance matrix for the regression is proportional to the inverse Hessian. Therefore the eigenvalues shall pack the fundamental information regarding fitting variance, which is essentially a non-negative quantity.

$$\arg\min_{\mathbf{p}_k}(\frac{1}{2}\mathbf{p}_k^T\mathbf{Q}\mathbf{p}_k + \mathbf{b}^T\mathbf{p}_k) \text{ subject to } \mathbf{a}_i^T\mathbf{p}_k = 0 \tag{3.2.24}$$

If this condition does not hold for some constraint outside the working set; we can still modify this expression by introducing a multiplicative weight factor to the optimal step which we have found from the optimization of the equality constrained solution;

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k \text{ with } \alpha_k \in [0, 1) \tag{3.2.25}$$

The additional factor can be identified in a systematic fashion if we account for the possible effects of iterative moves on the constraints which are excluded from our working set.

$$\mathbf{a}_i^T(\mathbf{x}_k + \alpha_k\mathbf{p}_k) \geq 0 \rightarrow \alpha_k \leq \frac{\mathbf{b}_i - \mathbf{a}_i^T\mathbf{x}_k}{\mathbf{a}_i^T\mathbf{p}_k} \tag{3.2.26}$$

$$\alpha_k \doteq \min(1, \min_{i \notin W_k, \mathbf{a}_i^T\mathbf{p}_k < 0} \frac{\mathbf{b}_i - \mathbf{a}_i^T\mathbf{x}_k}{\mathbf{a}_i^T\mathbf{p}_k}) \tag{3.2.27}$$

The results of this search is used to identify whether there are blocking constraints in the direction of our move or not. The decision condition for the update of working set thus can be reached through a two-case statement;

$$\alpha_k = \begin{cases} \alpha_k = 1 \text{ No blocking constraints, the move is feasible} \\ \alpha_k < 1 \text{ There is a blocking constraint outside of the working set} \end{cases} \tag{3.2.28}$$

A blocking constraint herein is identified as a violated inequality as it is a step outwards the feasible domain α_k.

3.2.1 Linear Constaints

1. **(Type 1) Stability of Excitations at Intermediate Compositions:** Stability of excitations are determined by their predicted positions relative to the tie line generated by the two closest tail compositions. The breaking of formation energy convex hull is prevented by introducing the pertinent linear constraint;

$$E_i - (\frac{c_i - c_A}{c_A - c_B}E_B + \frac{c_B - c_i}{c_A - c_B}E_A) \leq \delta E \tag{3.2.29}$$

This expression is nothing but the application of lever rule in phase diagrams. However, the pertinent expression that is employed in the quadratic programming entails the use of explicit expressions for the energy as a superposition of cluster functions;

$$\sum_{\alpha}^{\alpha_{max}} m_\alpha V_\alpha \left(\langle\sigma_\alpha\rangle_i - (\frac{c_i - c_A}{c_A - c_B}\langle\sigma_\alpha\rangle_i + \frac{c_B - c_i}{c_A - c_B}\langle\sigma_\alpha\rangle_i) \right) \leq \delta E \tag{3.2.30}$$

From a computational perspective it is desirable to link the concentrations $c_i, c_A \text{and} c_B$

to the point correlations. For the case of conventional spin assignments {1, -1} this correspondance can be written as;

$$c_i = \frac{1}{2}(\sum_p \langle \sigma_p^{point} \rangle_i - 1) \tag{3.2.31}$$

$$c_A = \frac{1}{2}(\sum_p \langle \sigma_p^{point} \rangle_A - 1) \tag{3.2.32}$$

$$c_B = \frac{1}{2}(\sum_p \langle \sigma_p^{point} \rangle_B - 1) \tag{3.2.33}$$

2. **(Type 2) Energy Brackets for Ground State Structures:** The second type of constraint is the introduction of upper and lower bounds for the DFT ground state energies. The motivation for this type immediately follows from the observation that a closer overlap between the predicted ground state energies and the DFT data shall result in an improved ground state line reproduction. The explicit expression is given as;

$$\delta E \leq E_i \leq \delta E \rightarrow \delta E \leq \sum_{\alpha}^{\alpha_{max}} m_\alpha V_\alpha \langle \sigma_\alpha \rangle_i \leq \delta E \tag{3.2.34}$$

3. **(Type 3) Stability of Excitations at Ground State Compositions:** The last two constraints are fundamentally of the same type. For the low energy excitations we first desire a correct oder of stability at a given composition as well as the correct energetic differences.

$$\Delta E_{DFT} \leq \Delta E_{CE} \leq \Delta E_{DFT} + \delta E \tag{3.2.35}$$

$$\Delta E_{DFT} \leq \sum_{\alpha}^{\alpha_{max}} m_\alpha V_\alpha (\langle \sigma_\alpha \rangle_i - \langle \sigma_\alpha \rangle_{GS}) \leq \Delta E_{DFT} + \delta E \tag{3.2.36}$$

The reason we introduce the 4^{th} type of constraint which solely sets the correct ordering of energies (for the sake of physical thermal behaviour) is the fact that imposing Type 3 constraints on all the phase space would yield over-stringent constaints for the quadratic programming problem. Therefore, as we move away from the convex hull, we become less stringent. This is equally valid for the intermediate compositions for which there is no ground state.

3.3 Fitting Metrics for the Expansion

Root mean square error (RMSE) analysis is the most conventional resort for assessing the goodness of a regressive fitting. The generic formula for the calculation of RMSE is given as[2];

$$RMSE(\hat{E})^2 = \frac{1}{N}\sum_{i=1}^{N}(E_i - \hat{E}_i)^2 \tag{3.3.1}$$

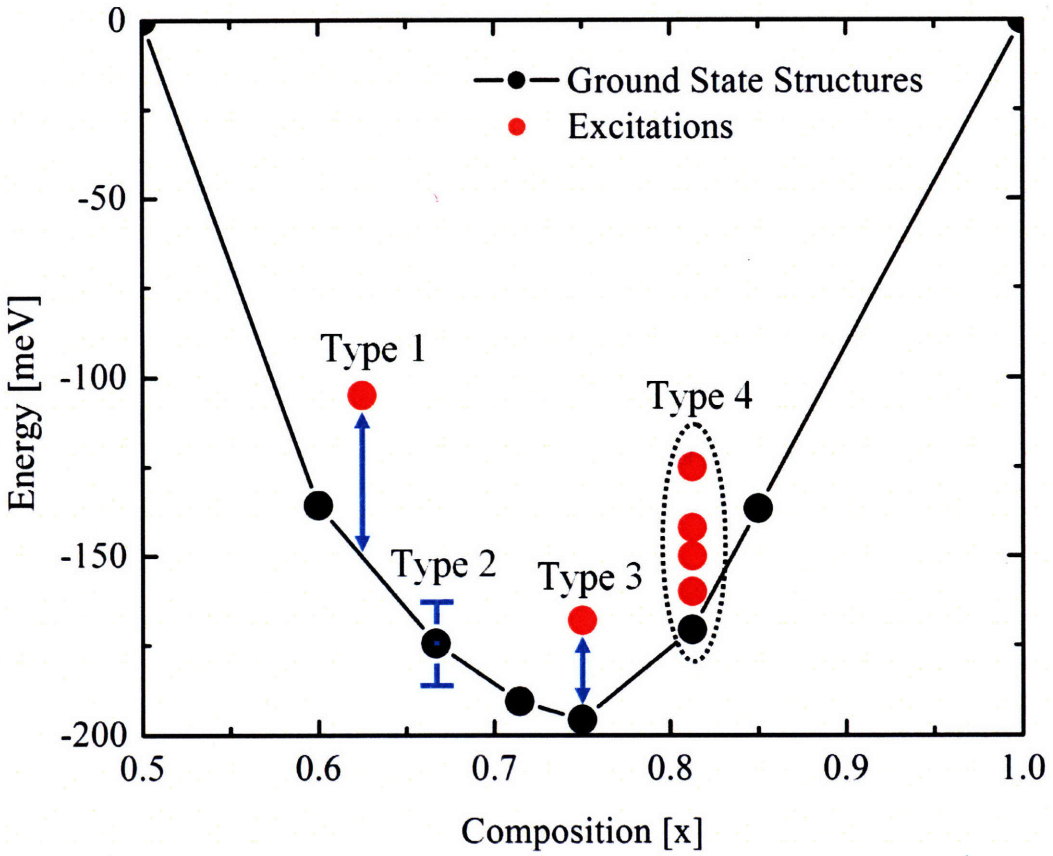

Figure 3.4: The illustration of basic linear constraints imposed on the cluster expansion

It is possible to show that RMSE can be decomposed into the sum of variance and bias pertaining to the estimator[33]. There is usually a trade-off between these two properties and we try to strike a balance in fitting the data as neither high bias nor variance is desirable. Owing to its nature, a lower RMSE will indicate a closer overlap (in the least squares sense) between the data points and their predicted values. As we add more terms to the expansion, the predictions at each point will progressively get closer to their actual values. If we track this behavior, the RMSE will be a monotonically decreasing function of added terms. However, this comes with price: the danger of overfitting. This regressive problem caused by the use of excess free parameters in the expansion. The additional degree of freedom manifest itself as a fitted function faithfully trailing noise. Although the apparent RMSE value stays very

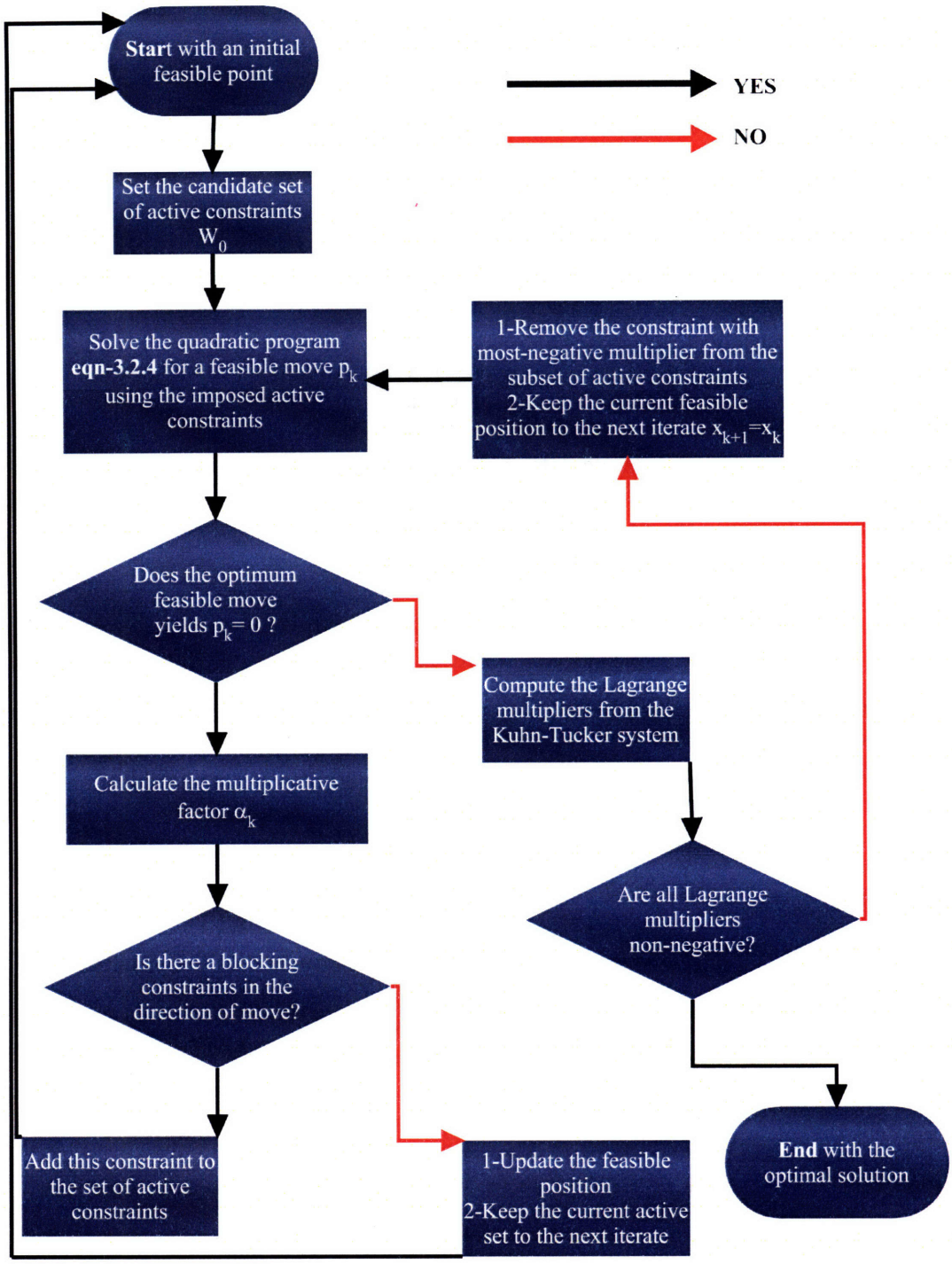

Figure 3.5: The Flow of Active Set Algorithm for Quadratic Programming with Linear Inequality (and Equality) Constraints

low, the prediction of additional points will suffer from the unsmooth behavior in the rest of

sample space.

Cross Validation(CV) is another procedure by which one assesses the goodness of the fit[1, 12]. The advantage over the RMSE in ordinary least squares is the extra information regarding the predictive power of regression. The generic formula for CV score is given as;

$$CV^2 = \frac{1}{N}\sum_{i=1}^{N}(E_i - \hat{E}_{(i)})^2 \tag{3.3.2}$$

The E_i represents the energies calculated from an external source such as the DFT calculations whereas $\hat{E}_{(i)}$ is the energy predicted from the least squares regression computed with the remaining $N-1$ structures.

The generic formula, albeit being intuitive is not a computationally efficient way of calculating the CV score as it stands. As it entails us to carry out a regression for each point left of; the computational complexity will scale with a second order polynomial N^2. There is, however, a more compact way of calculating the CV score which exploits a single least squares regression with all the structures included. As it assumes linear complexity the alternative formulation below is the commonly adopted means to calculate the leave-one-out CV score[33].

$$CV^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{(E_i - \hat{E}_i)^2}{1 - \mathbf{x}_i(\mathbf{X^T X})^{-1}\mathbf{x}_i^T} \tag{3.3.3}$$

X is the matrix of correlations with dimensions $N \times M$ where N is the total number of structures and M is the number of clusters included in the least squares fit. The qualitative behavior of a CV graph can be understood by the competition of two terms in the numerator and the denominator. The numerator term from the ordinary least squares steadily decreases with increasing M and analogous to the RMSE error. As we add more clusters this term will tend to bring down the CV score. The denominator, however, requires more subtle analysis. We start by looking at the expression $x_i(\mathbf{X^T X})^{-1}x_i^T$. If the x_i vectors were orthogonal this would yield an element of the trace of $(\mathbf{X^T X})^{-1}$ matrix. The trace sum itself is proportional to the variance of the least squares estimator. As the overfitting sets in, the variance increases since the noise content is more dominant. Therefore the eigenvalues of the covariance matrix will be more evenly distributed rendering the denominator to approach zero. Such behavior will more than compensate the plummeting difference between the input values and the predictions.

The behavior of the RMSE and the CV score in a real cluster expansion (with no constraints or weighting imposed) is illustrated via the graph below. The RMSE is indeed a monotonically decreasing function of added terms whereas CV score soars up rapidly with the onset of overfitting.

It is also important to note that the actual evolution of CV score shall not be a strictly U shaped convex function as there is no unique way of imposing an order in the inclusion of clusters to the system at hand.

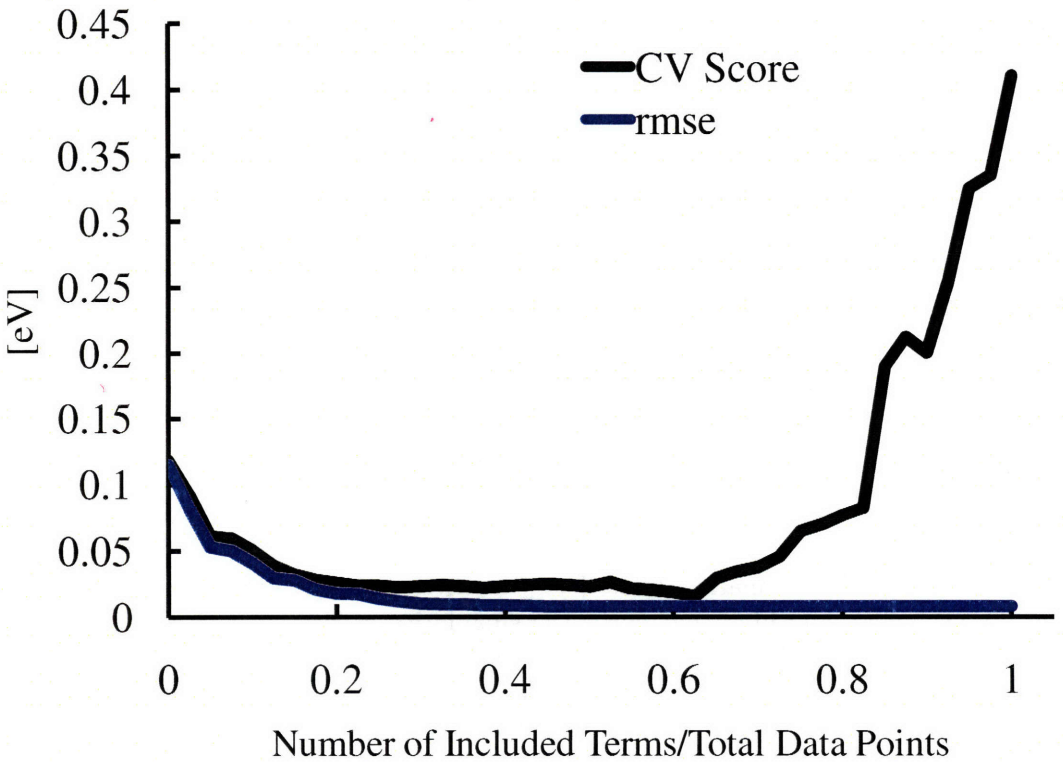

Figure 3.6: Comparison of the generic behavior of the RMSE and CV Score as the number of the terms included in the expansion is increased

3.4 Regularization with Kernel Smoothing

Regularization of Ill-posed problems were first introduced for solving certain families of integral equations with square integrable kernels hence coining the term kernel smoothing[16, 35]. The most famous of these are the Fredholm equations of the first kind whose solutions can be perturbed to have arbitrarily large norms following small perturbations in the transformed domain through the high frequency terms. Therefore, the very definition of ill-posedness assumes the property of infinite dimensionality. However; it is still possible to find discrete problems which are not merely a discretization of a continuous counterpart. Hansen lists two criteria for characterizing a discrete ill-posed problem[16];

1. The singular values of the design matrix **A** should decay to zero with no particular gap (or series of gaps) in the spectrum. This rules out the possibility of finding a well-posed problem close enough to the original problem.

2. The conditioning number based on singular values, i.e. the ratio of the largest singular value to the lowest singular value, shall be large. This is the quintessential symptom of

an ill-posed problem. The larger the conditioning gets the more sensitive the solution will be to the perturbations.

Regularization with Kernel methods is a vast subject with rigorous functional analytic foundation. Yet, (luckily in an elegant fashion) its implications can be formulated and incorporated easily. The main aim in regularization is to gain freedom for incorporating as many terms as we like to the cluster expansion. In the conventional least squares procedure, this is a challenge plauged by the singularity of correlation matrices.

The way regularization improves the ordinary least squares procedure is most aptly illustrated by looking at the modified formulation of the original quadratic problem and the resulting minimizer;

$$\arg\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}_{DFT}\|_2 + \lambda^2 \|\mathbf{Lx}\|_2 \tag{3.4.1}$$

$$\hat{\mathbf{x}} = \mathbf{A}^T(\mathbf{A}^T\mathbf{A} + \lambda^2\mathbf{I})^{-1}\mathbf{Ax} \tag{3.4.2}$$

Adding a multiple of the identity matrix to the Hessian shall counteract the lowest singular values of the Hessian matrix thus providing a solution for the ill-conditioning. Remembering that the covariance matrix differs from the inverse Hessian upto a multiplicative factor (the RMSE error), anything that curbs the singularity of Hessian will automatically reduce the variance of regression. The method we propose for the regularization of a constrained quadratic program is to feed this modified Hessian to the active set algorithm we have described above. The main idea can be illustrated by reformulation of the quadratic programming problem;

$$\arg\min_x(\frac{1}{2}\mathbf{x}^T(\mathbf{A}^T\mathbf{A} + \lambda^2\mathbf{I})\mathbf{x} + \mathbf{c}^T\mathbf{x}) \text{ subject to } \begin{cases} \mathbf{D}^{\mathbf{T}}_{\mathbf{ineq}}\mathbf{x} \leq \mathbf{f}_{\mathbf{ineq}} \\ \mathbf{D}^T_{eq}\mathbf{x} = \mathbf{f}_{eq} \end{cases} \tag{3.4.3}$$

The main ingredient of the strategy described herein is probing of the optimum value of regularization parameter. The two criterions we shall be describing will be the Generalized Cross Validation score (GCV) and the L-Curve criterion. The generalized cross validation score is a generalization of the ordinary cross validation statistics of Allen the main difference being the rotational invariance[1, 13]. Such an invariance can be argued to be a characteristic of a good estimate for λ when the prior distribution has a spherically symmetric density for the ridge estimate and the error vector pertaining to the model. It is defined as;

$$\text{GCV} = \frac{\|\mathbf{Ax}_{regular} - \mathbf{b}\|_2^2}{(\mathbf{trace}(\mathbf{I_m} - \mathbf{AA^I}))^2} \tag{3.4.4}$$

with the explicit formulation of the matrix A^I as;

$$A^I = (\mathbf{A}^T\mathbf{A} + \lambda^2\mathbf{I})^{-1}\mathbf{A} \tag{3.4.5}$$

The final formulation, when compared to the conventional definition of cross-validation bears two differences which are namely the presence of trace in the denominator(bringing out the rotation invariance) and the modified Hessian. It has been shown that the value of ridge parameter minimizing this expression shall naturally be the optimum regularization

parameter choice[13]. However, in numerious accounts it turns out that a choice based on GCV is too small[17]. However, when it works GCV provides a very good estimate for λ.

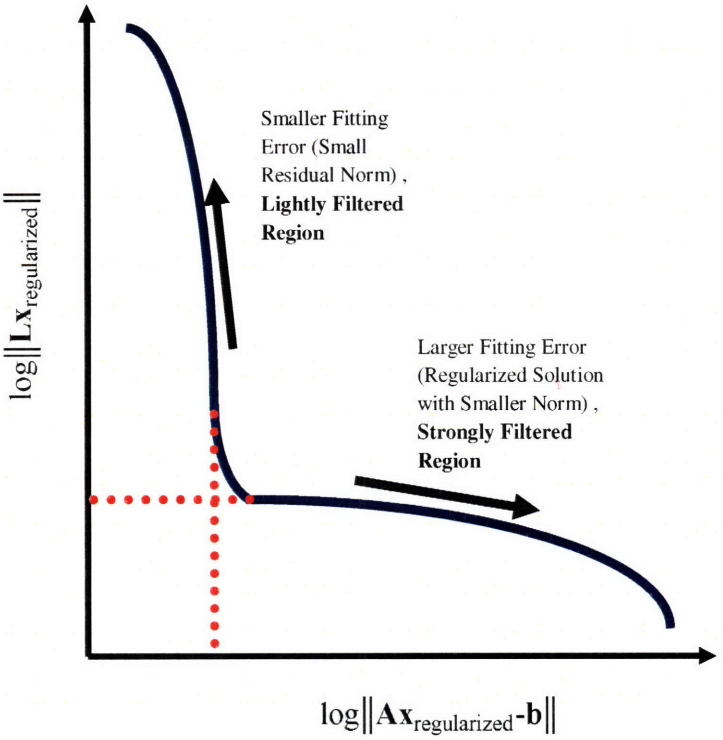

Figure 3.7: A continuous L-curve with two distinctive regions for a regularized problem[16]

The other method for computing the optimal Tikhonov parameter is the L-curve method. This method takes advantage of a very intiutive assumption that the optimum parameter choice should seek a balance between a good least squares solution with small fitting error, and a small norm for the regularized solution, i.e impeding a possible ECI blow-up. The point yielding the best trade-off between these two factors than can be used to calculate the desired ridge parameter. The construction of such a curve is shown in **fig-3.7**. This method is the adopted choice for regularization in this work while the concept of generalized cross validation is used as the fitting metric for subset selection.

The L-curve has a generic shape (literally resembling the letter L, hence the name) marked by two regions which represent different amounts of regulative filtering. The clearly visible corner marks the point where we calculate the ridge parameter. This can be most naturally done by spotting the curvature of the L-curve at that point. Classical differential geometry endows us with the closed formulas to calculate curvature at a given point (as it is defined at a point). However, in a numerical run, we do not retain the luxury of sampling a continuous range of ridge parameter choices. To overcome this we operate on a discrete set of parameter choices and connect them to form a continuous L-curve with splines[16, 17]. Then we can spot the corner of the curve which will yield the optimum choice among different options. As the shape of this corner is noticeably sharp we shall spot a spike in the graph pertaining to the curvature as a function of the ridge parameter. However, as we iterate over many steps (for each candidate cluster set) we do not explicitly draw the corresponding L-curve but merely spot critical points at each iterate and feed into the consequent quadratic program.

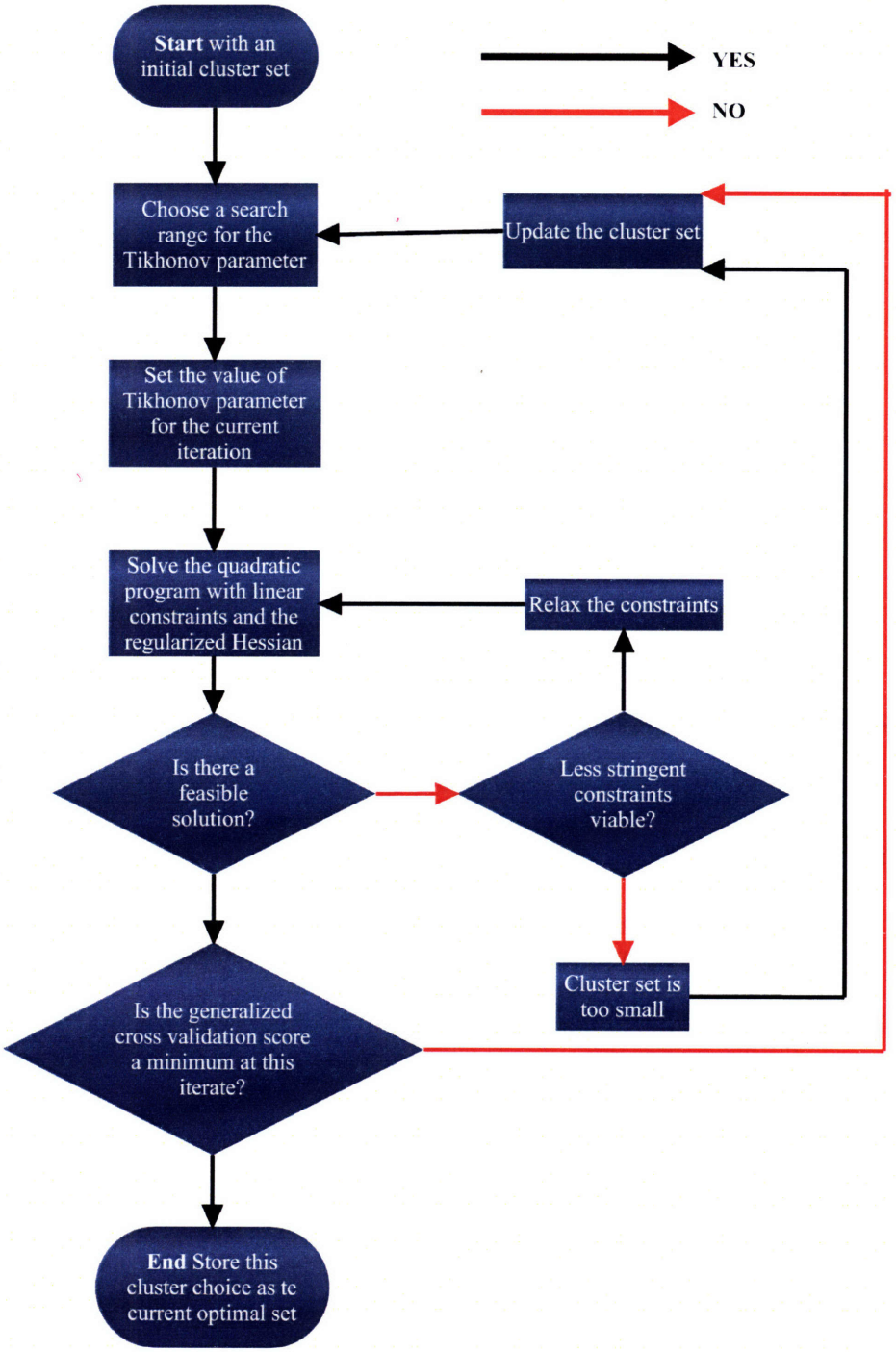

Figure 3.8: General cluster expansion algorithm with kernel smoothing and quadratic programming

3.5 Possible Extensions and Additional Exploratory Tools

3.5.1 Order Parameter Loops and the Dynamic Update of Interactions

Order parameters are vital tools in probing phase transitions. They have been introduced by Lev Landau with the observation that phase changes are almost always accompanied by a pertinent symmetry breaking[9, 22]. According to their nature and underlying physical phenomena they can be scalars, vectors or tensorial quantities defined over the real or complex fields (in the mathematical sense). As general as their mathematical nature, their definition also bears little constraints on physicists side. An order parameter thus a quantity we define as long as we have a proper understanding of the broken symmetry.

For the case of Ising models, the natural order parameter is a scalar (the concentration difference, magnetization and so on). In our particular case of Na_xCoO_2 it can be defined as the difference between occupancies on Na(1) and Na(2) sites. In a system like Na_xCoO_2 for which electrostatics and site energy differences are at a competing play, the proper choice of relative strenghs of these terms can be crucial. The current method can be concurrently used with Monte-Carlo simulations to explore the relative strength of pair interactions and site energy differences so that the cluster expansion will not yield ground states which are phase seperated or, as the opposite case, a frustrated supersized lattice when the repulsive interactions dominate the Hamiltonian only to be counterbalanced by line defects on the Na lattice.

The algorithm defined in this work provides a robust and modular tool for exploring competing systems as we can both dictate the absolute magnitude and the sign of individual interactions and their relative strengths. Systematic sweeping of these properties brings a significant exploratory freedom in understanding of the physics. In this way it is literally possible to teach physics to the cluster expansion.

3.5.2 Structure Selection

Structure selection has been introduced formerly for ruling out outliers from the input set of DFT energies[37]. The algorithm essentially works by choosing a subset of input structures so as to minimize the variance of fit. Therefore at each step we seek to obtain maximum variance reduction by removing a structure from our initial working set. Remembering that the covariance of a least squares fit is given by the following trace sum;

$$VAR = \mathbf{trace}(RMSE^2 \mathbf{A}^T \mathbf{A})^{-1} \tag{3.5.1}$$

We seek the smallest eigenvalue of the matrix $\mathbf{A}^T\mathbf{A}$ and the corresponding eigenvector as that is the direction of search for selecting structures. This can be most easily included into our expansion route as a preprocessing tool, i.e. before we adjust our ECI in accordance with the imposed constraints. Structure selection, as useful as it is, is not a tool we should resort in the first place whenever we embark on a cluster expansion. As we do not have any knowledge regarding the nature of errors in our input energies, avoiding structures for reducing variance may introduce strong bias to our fit and might give wrong disordering energy and shift the temperature scale in Monte-Carlo runs.

Chapter 4

Cluster Expansion on a Model System

We apply the quadratic programming technique to the model system of Na_xCoO_2 in the composition regime x=[0.75, 1]. Na_xCoO_2 is a complex transition oxide of continued scientific interest owing to its exhibition of a gamut of unique physical behaviour as Na atoms are removed from the stoichiometric compound[34, 32, 30]. The high mobility of ions stemming from the tunneling through vacant sites of Na sublattice is identified as the main source of collective phenomena for this material[19, 28]. Those properties span a broad range such as the superconductivity of the hydrated form at around x=0.11, and the high Seebeck coefficient at around x=0.75. As a mixed valence oxide, it is close to the electronic phase behaviour including the spin transition or the metal insulator transition.

The closest oxide system Li_xCoO_2 is the primary cathode material in rechargebale batteries and has been shown to display rich compositional and electronic phase behaviour. The most evident common property is the octahedral complex of cobalt d-orbitals and the surrounding oxygens. The amount of Na is also key to the valence state occupancy of the Co layers. It is the first reported example of a first order Mott (Metal Insulator) transition following Li deintercalation at room temperature[24]. All the known examples from doped crystalline materials exhibit continuous phase changes despite the original prediction of a discontinuous transition between the metallic and insulating states of impurities. The high mobility of Li ions is thought to be the key to the discontinuous nature as it allows impurity ordering. Since the impurities are not configured randomly, the possibility of an Anderson localization is practically ruled out thus corroborating the exploratory interest in the analogous system Na_xCoO_2. In this respect, $Na_xCoO_2$is a promising playground for assessing the basic understanding of transition metal oxides and mixed valence systems. The electronic phase diagram proposed for this system, albeit being a primary source of physical knowledge inherently lacks the details of atomic ordering and the complete justification of physics driving it[19].

On the experimental side there are numerous experimental methods of Na removal from the compound which can be categorized into processes of chemical or electrochemical nature. Combined with the electron and neutron diffraction experiments they constitute the fundamental tools providing vital evidence for the nature of atomic ordering in nonstoichiometric compounds. Relying on such procedures, there have been numerous studies published on the Na ordering in the Na_xCoO_2, yet there has not been a complete phase diagram (in the space of Na composition) hitherto[40, 25, 19, 28]. There is no consensus over the ground states

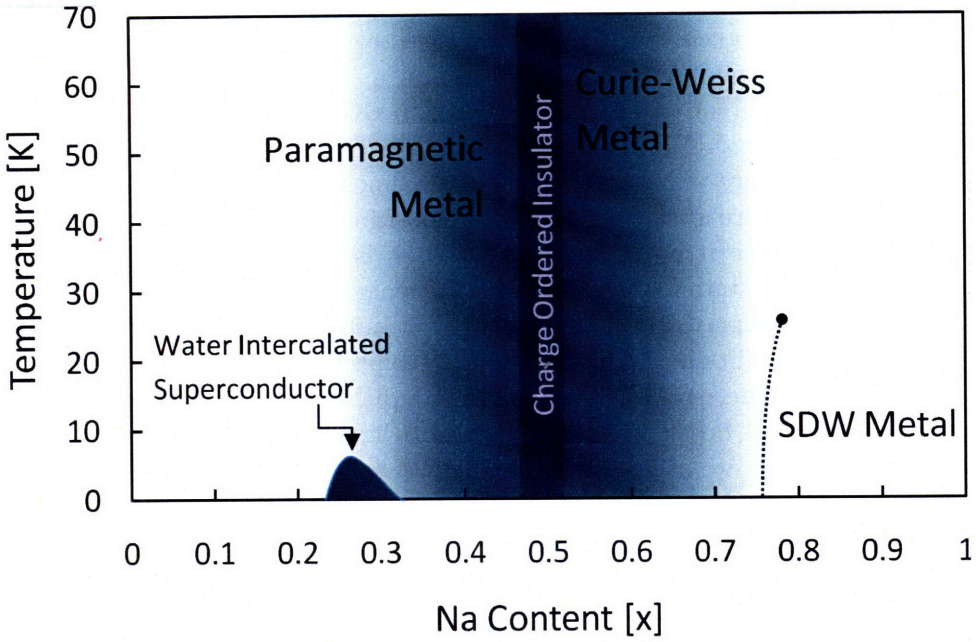

Figure 4.1: Proposed electronic phase diagram of Na_x CoO_2

on different compositions. The values pertaining to the ground state compositions have not been unambigiously established either.

Among possible stacking alternatives, the structures explored here pertains to the class P2 which assumes an AABB type stacking of oxygen layers along the c axis. This particular structure forms the parent lattice for the cluster expansion we present in this work.

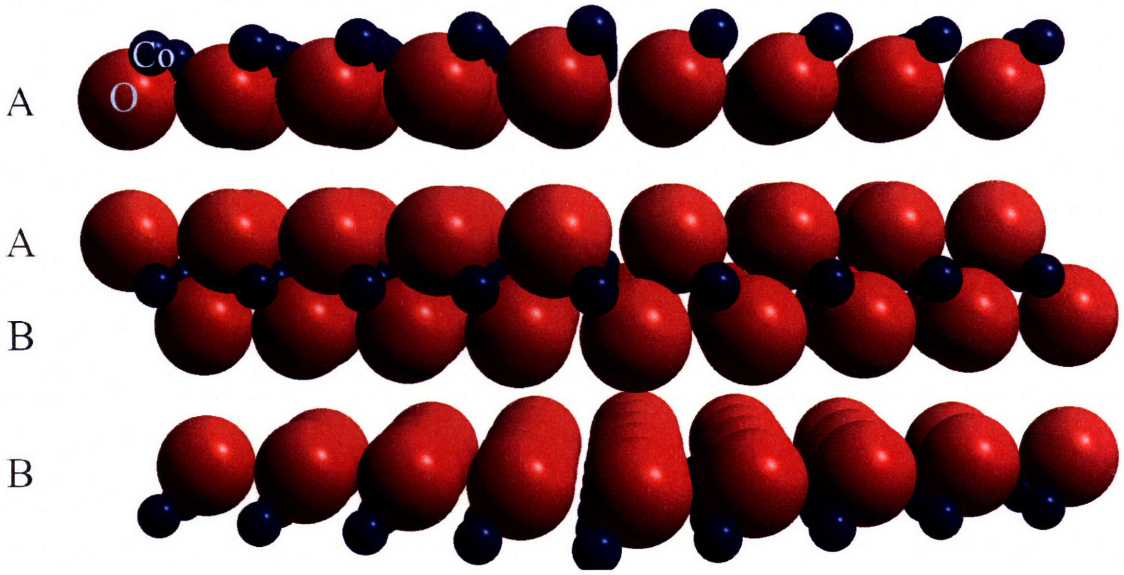

Figure 4.2: P2 stacking in $NaCoO_2$

The composition range x=[0.75, 1] is a relatively safe regime for relying on the GGA methodology as the explicit charge ordering and the coupling from the Co sublattice will be

less pronounced. It can thus be argued that the nature of electronic effects shall be uniform in accordance with the electronic phase diagram.

The input set comprise 39 different structures calculated in GGA, and form a convex hull with 4 identifiable ground states at x=0.75, x=0.8125, x=0.85 and x=1.0. As the consequitive slope changes in the convex hull are very subtle(yielding an almost degenerate ground state line) and the lowest energy structure at the metastable composition 0.8 is very close to the tie line generated by the ground states at x=0.75 and x=0.8125, it is very critical that a cluster expansion reproduces the ground state line with perfect agreement. As all the mathematical routes for regression minimize the pertinent fitting metrics for uniform sampling of the phase space, tuning of the ECI with additional tools is inevitably expected to bring the values of metrics up. However, the complete recovery of the convex hull has been long known to be crucial in getting the correct phase boundaries amd thermal behaviour (relative stability and the temperature scale). Therefore, we utilize constrained quadratic programming with linear constraints concurrently with regularization and target at the full recovery of the ground state line rather than seeking the absolute minimum of a fitting metric.

The candidate cluster list comprise the empty term, 2 possible point terms , as well as the pair and triplet terms upto a limiting size of 9 angstroms on the parent lattice. Site occupancy by Na is denoted by the spin assignment 1 and the Na vacancy is represented by -1. The existence of two point terms stems from the fact that there are two Na sublattices which can be distinguished with symmetry in our parent lattice. These are denoted as as Na(1) and Na(2) and they are illustrated below;

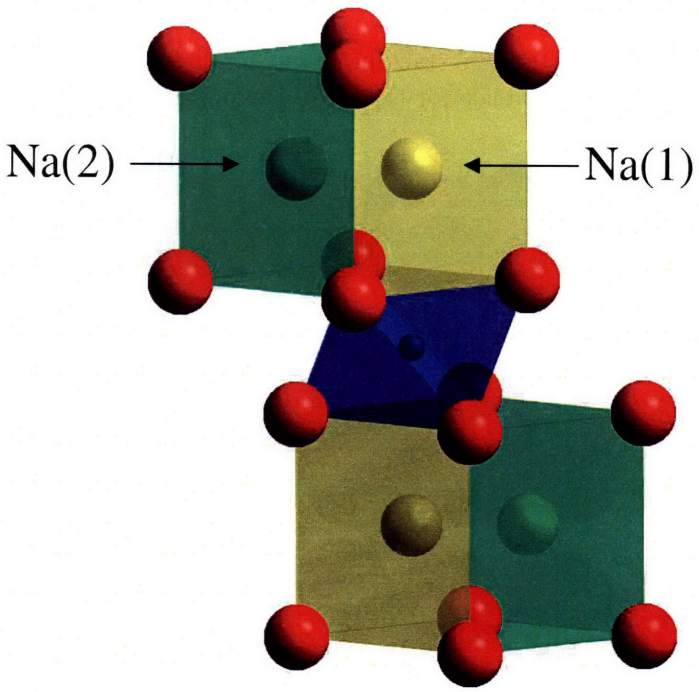

Figure 4.3: The atomic structure of Na_2CoO_2 showing the two possible Na sites.

The 29 ECI included in the expansion are grouped in size and respective interaction type, and plotted against distance below;

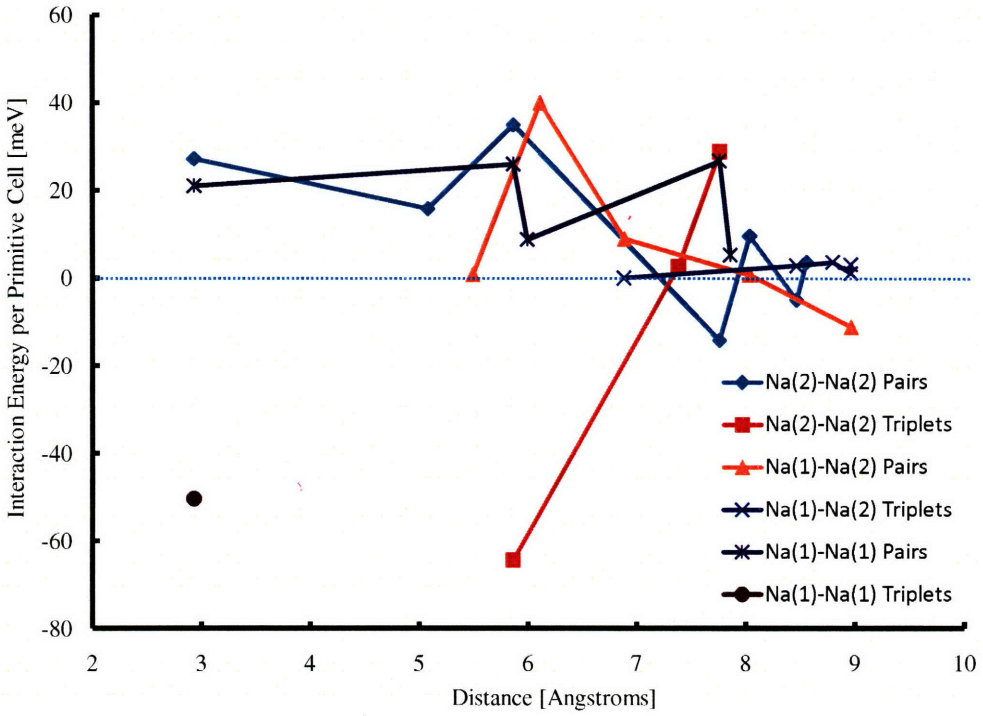

Figure 4.4: Distance dependence of effective cluster interactions

In the expansion presented herein the site energy difference between two sites is around 90 meV per formula unit. This shows that there is a noticeable preference for Na(2), when our sole concern is introducing a new Na to the repulsive mean field of the Na sublattice. Looking closely at the remaining interactions we observe that the first nearest neighbour pair interactions in both sublattices are repulsive and there is a general repulsive trend for pair terms. The most noticeable negative terms in the expansion are the nearest neighbour 3 body terms and they denote the interaction of a Na(1) (Na(2)) site with a neighoring nearest neighbour Na(1)-Na(1)(Na(2)-Na(2)) pair. This particular behaviour interesting as it will favor filled neighbouring sites of the same site whenever we have fully occupied nearest neighbour pairs and corroborates the view that the GGA picture is a manisfestation of site energy differences competing with electrostatics in seeking a balance between an increased average Na distance and Na(1) site occupancy (as it is expected and found to be more favourable to occupy)[28].

The expanded convex hull reproduce the ground state line in full acordance with the GGA calculations. The ordering of lower energy excitations upto 20 meV (set through the quadratic program) from the convex hull are also recovered with the help of linear constraints imposing the recovery of GGA predictions. The tuning of the ECIs at will is partly possible with the large number of clusters (each bringing an additional degree of freedom in the Hamiltonian) in the expansion. However, the freedom of cluster choice and using an arbitrary(almost) interaction set size is not possible in the conventional Connolly-Williams scheme of structure inversion with least squares[8]. This is inherently a numerical stability problem of mathematical nature yet it is entirely possible that it will have grave implications in the imitation of the accurate physical principles. Luckily this can be addressed with the tools of regularization which we have outlined in the former sections.

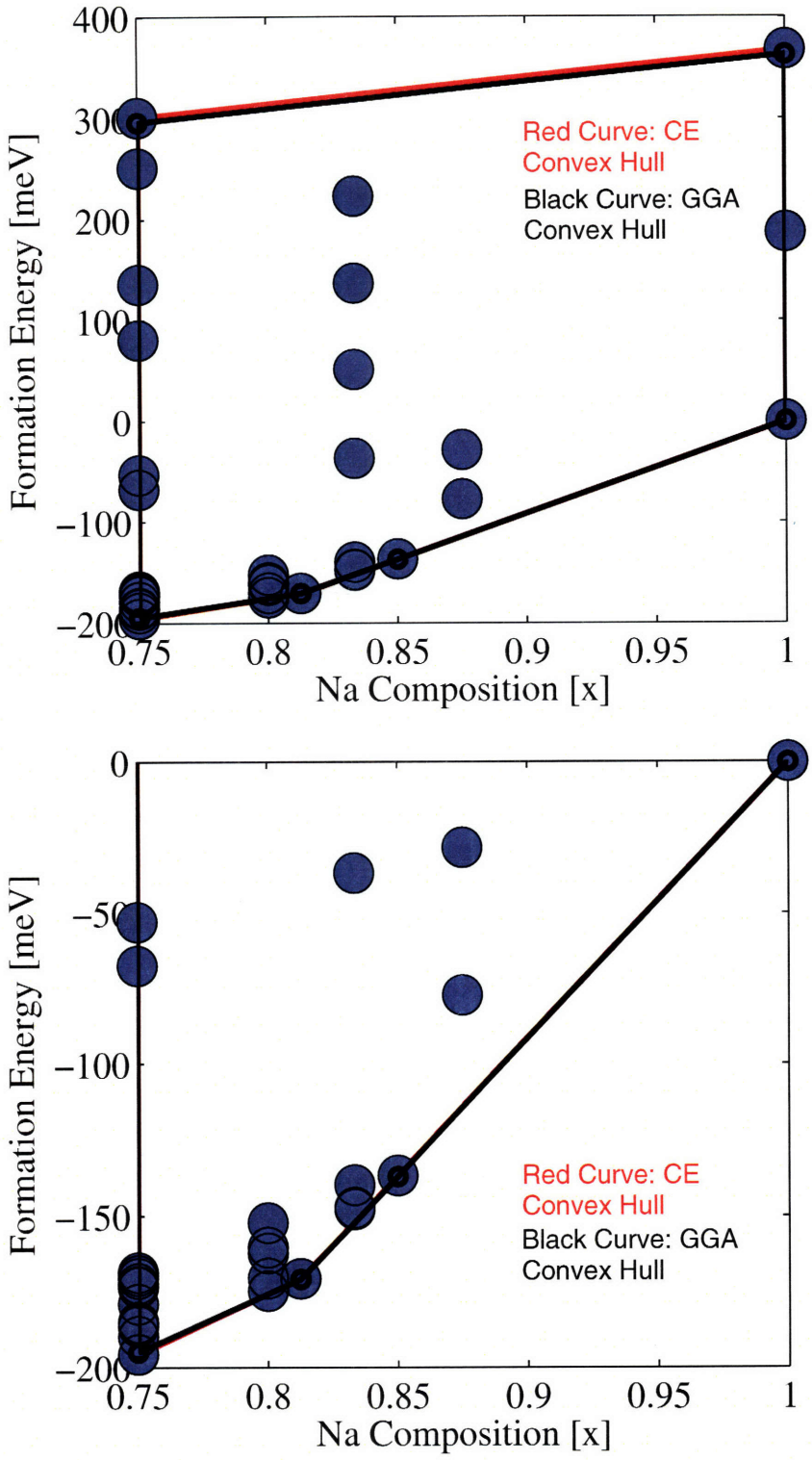

Figure 4.5: GGA convex hull superimposed on the cluster expanded hull

The decision to regularize a fitting problem starts with an unambigious idenfication of the ill-posedness of the design matrix. One of the tools described formerly is the SVD and it allows us to characterize the ill-posedness and numerical instability lurking in our system. Therefore before regularizing our system we first check the singular value decomposition to see whether we can spot noticeable gaps in the spectrum, especially in the lower magnitude tail.

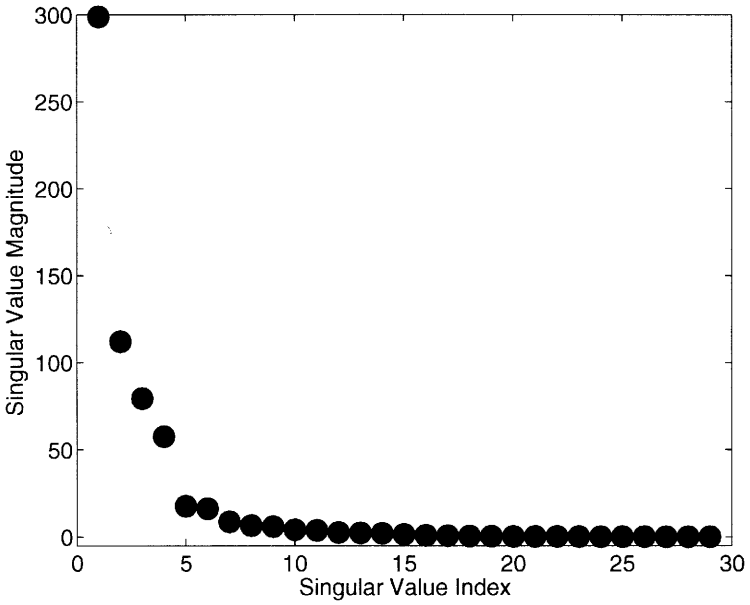

Figure 4.6: Decay of singular values of the Hessian

The plot of singular values indexed by respective magnitudes clearly shows that the spectrum is dominated by singular values of smaller magnitude therefore it is not possible to solve this problem by truncating the design matrix (by removing a few clusters from the interaction set). This basically means that there is no nearby expansion with a non-singular correlation matrix which would conveniently replace the original matrix.

There is another crucial symptom regarding the singular value decomposition of the Hessian which is characterized by rampant sign changes in the singular vector elements of each basis vector. Although there is no rigorous mathematical proof so far, frequent sign changes have strong implications on a possible ECI blow-up which is unaccaptable as each and every member of the ECI set has a clear physical assignment. Albeit the absence of a proof, we can still corroborate this argument by the simple implications of sign changes. If we think of the ECI we can see that the cluster expanded quantity is calculated by adding all contributions from different cluster interactions. Therefore a small bounded value can either be accomplished by summing up small individual terms or by arbitrarily large contributions which are counterbalanced by other interactions of opposite sign and still resulting a bounded energy. This also recapitulates the whole idea of regularization in impeding the largely overshoot ECI values and a prohitibively large solution norm.

If we look at the relative ratio of the number of elements of different algebraic sign we see that there is a clear trend of noticeable balance between the counts of each sign for each vector. Therefore we might expect that a matrix of this nature will result in unphysical ECIs

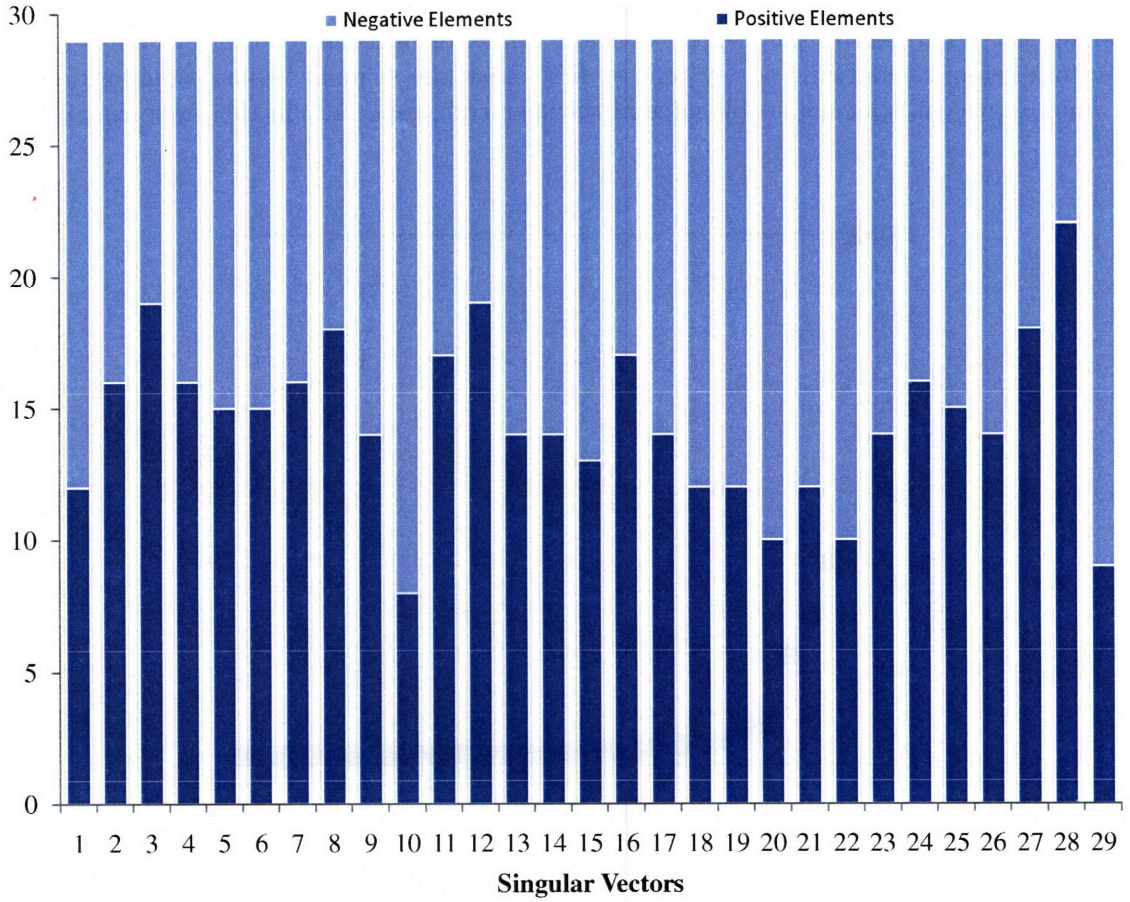

Figure 4.7: Sign-distribution of singular vector elements

even if the predictions and the metrics of expansion does not show a clear sign of instability and ill-posedness. The succession of alternating signs thus provides another perspective for the danger of blown-up interactions and the need for regularization.

The regularization procedure, as its algorithmic layout suggests, is an iterated procedure, i.e. for each subset of interactions (corresponding to a different sized matrix with distinct numerical conditioning) we have to calculate an optimal Tikhonov parameter. This naturally entails that we construct an L-curve for each iteration and spot the corner marking the maximum value of curvature. After the optimal value is obtained; it is used to regularize the Hessian matrix in our quadratic program. In a way, we tweak the curvature of objective landscape before starting the optimal solution search through the boundary of constraint functions via the active set route. Therefore, at each iteration we have to choose a suitable parameter in accordance with the system dimension (size of the cluster set) and the underlying topography (Hessian spectrum). These two properties shall also determine the detailed shape and curvature traits of the L-curve albeit the generic shape shall always be retained thanks to the general arguments of regularization.

The particular L-curve for the final cluster expansion with 29 terms indeed has the generic

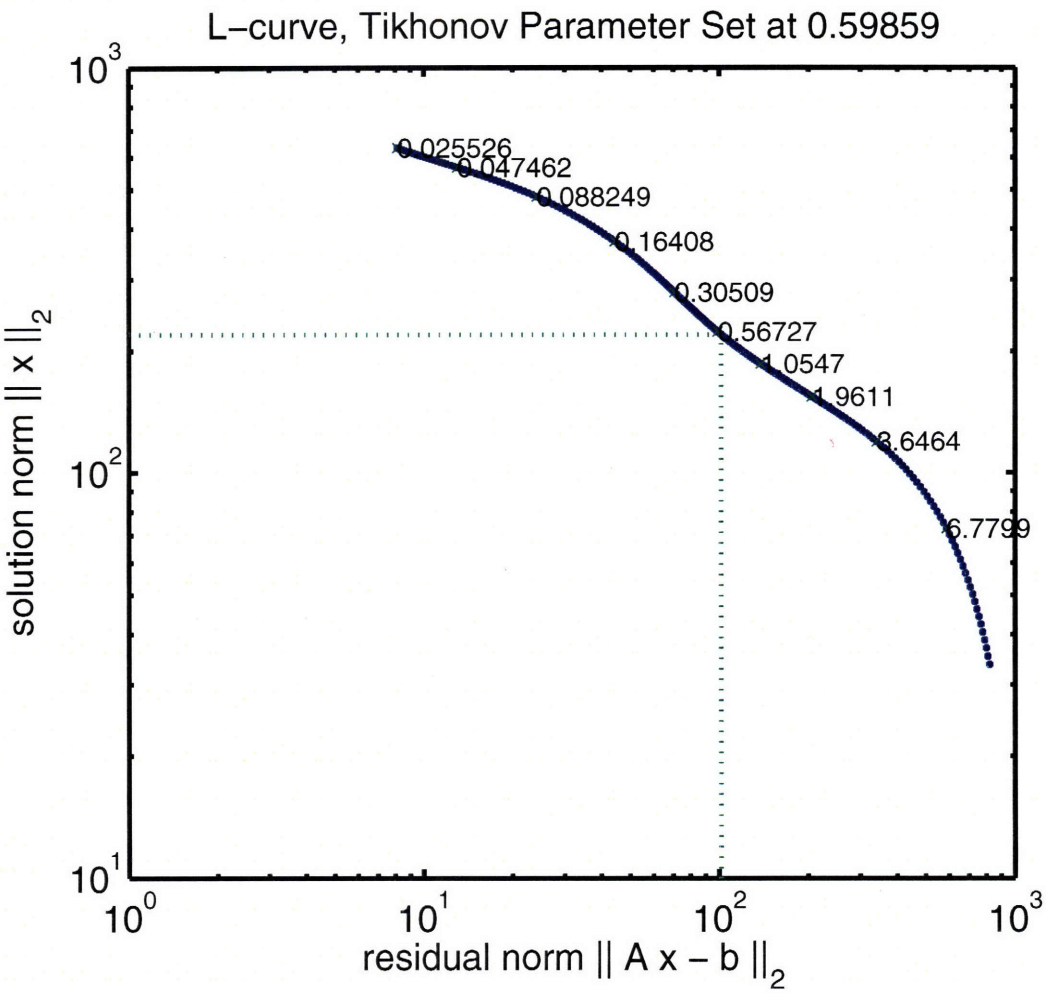

Figure 4.8: L-curve for the final set of clusters

shape but the curvature traits are rather broadened with an optimal parameter estimate at $\lambda = 0.59859$. This means that the ill-posedness, being evident from the SVD analysis can also be remedied with utilizing a safer (broader) range for parameter estimation. In other words, the trade-off between the solution norm and the residual norm is weakly manifested for this particular problem. As the amount of filtering imposed does not lead to steep movements on the L-curve; we could have switched between alternative methods of parameter estimation with less surprises on the numerical side.

Chapter 5

Final Discussion

The general scheme of merging regularization with quadratic programming works by culling different methods introduced so far into a single robust package. This gives the strength of using concurrent tools without compromising the control over fitting parameters. The main idea is to retain the physical control over our system rather than being guided with a tool which might be severly limited by the underlying numerics. The fundamental idea of constraints, all of which are choosen to have clear physical significances allow us to actually teach physics to the the automated algorithm. Instead of relying to the weighting of low energy domain of the phase space in recovering the ground state line, we can confine our search to the region of phase space in accordance by the rules we define and impose. However, the modular nature and the generality of the underlying assumptions and tools allows us for continual upgrading and exploration with state of the art methods. Moreover, using the regularization tools simultaneously frees ourselves from the worries of singular design matrices that might pop up as we add more clusters. Therefore, the expansion scheme introduced in this work is more of a physical route than a mathematical one. This is however, both it's strength and it's weakness; as we gain more control over the physics there will be a stronger need for the proper understanding of the underlying physical principles so that we can take advantage of the its full scope. This brings out a conclusion which is already uttered by the first principles community; there is no way of fully automating phase diagram calculations without scrutinizing the ground states coming from the DFT calculations. The strength and the utility of the output will utterly determined by the amount of control and comprehension over the input we have in the physics of our systems of interest. Therefore, what we present here is not a panacea for automating phase diagrams but an intermediate tool for exploring properties of the complex lattice systems like the Na_xCoO_2 and the extraction of approximated energetics from first principles data.

A possible sequel to this work would be the determination of optimal choice of fitting metrics and regularization tools. That is, not just minimizing a given metric but first establishing what to minimize so as to define an optimal cluster expansion. Although seemingly a task that should be tackled by the stasiticians; the materials science community is the main correspondent as none of the fitting metrics employed herein or elsewhere are universal. The debate over free-lunches centered on the concept of cross validation makes it worthy if not obligatory to rigorously establish whether there is really a free lunch for case of cluster expansions[38, 14]. Yet another crucial question that awaits an answer is deciding if we can use physical principles (which is absent from most problems of pure mathematical interest) to ensure that the relevant fitting metrics will be properly functional for our final goal of

deciding whether we have an optimal expansion or not.

Bibliography

[1] David M. Allen. The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, 16:125–127.

[2] Roger J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences (Manchester Physics Series).* John Wiley and Sons, 1993.

[3] Michel Le Bellac, Fabrice Mortessagne, and G. George Batrouni. *Equilibrium and Non-Equilibrium Statistical Thermodynamics.* Cambridge University Press, 2004.

[4] Nicolas Bourbaki. *Elements of Mathematics Algebra I Chapters 1-3.* Springer-Verlag, 1998.

[5] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambrdige University Press, 2004.

[6] Christopher J. Bradley and Arthur P. Cracknell. *Mathematical Theory of Symmetry in Solids.* Clarendon Press, 1972.

[7] Stephen G. Brush. History of the lenz-ising model. *Reviews of Modern Physics*, 39:883 – 893, 1967.

[8] J.W.D. Connolly and A. R. Williams. Density-functional theory applied to phase transformations in transition-metal alloys. *Physical Review B*, 27:5169, 1983.

[9] Brian Cowan. *Topics in Statistical Mechanics.* Imperial College Press, 2005.

[10] James W. Demmel. *Numerical linear algebra.* SIAM Publications, 1997.

[11] Roger Fletcher. *Practical Methods of Optimization.* John Wiley and Sons, 2000.

[12] Seymour Geisser. The predictive sample reuse method with applications. *Journal of the American Statistical Association*, 70(350):320–328, 1975.

[13] Gene H. Golub, Michael Heath, and Grace Wahba. Generalized cross validation as a method for choosing a good ridge parameter. *Technometrics*, 21, 1979.

[14] Cyril Goutte. Note on free lunches and cross validation. *Neural Computation*, 9:1245–1249, 1997.

[15] Denis Gratias, Juan M. Sanchez, and Didier De Fontaine. Application of group theory to the calculation of the configurational entropy in the cluster variation method. *Physica A*, 113A:315–337, 1982.

[16] Per Christian Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion (Monographs on Mathematical Modeling and Computation).* SIAM Publications, 1987.

[17] Per Christian Hansen. Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34:561–580, 1992.

[18] Kerson Huang. *Statistical Mechanics.* John Wiley and Sons, 1987.

[19] Qing Huang, Maw Lin Fo, Jr Robert A. Pascal, Jeffrey W. Lynn, Brian H. Toby, Tao He, Henny W. Zandbergen, and Robert J. Cava. Coupling between electronic and structural degrees of freedom in the triangular lattice conductor Na_xCoO_2. *Physical Review B*, 70:184110, 2004.

[20] Sorin Istrail. Statistical mechanics, three-dimensionality and np-completeness i. universality of intractability for the partition function of the ising model across non-planar lattices. Proceedings of the 32nd ACM Symposium on the Theory of Computing (STOC00), Portland,Oregon, 2001. ACM Press.

[21] Erwin Kreyszig. *Introductory Functional Analysis with Applications.* John Wiley and Sons, 1989.

[22] Lev D. Landau and Evgeny Lifshitz. *Statistical Physics (Course of Theoretical Physics, Volume 5).* Pergamon Press, 1984.

[23] Melvin Lax. *Symmetry Principles in Solid State and Molecular Physics.* Dover Publications, 1974.

[24] Chris A. Marianetti, Gabriel Kotliar, and Gerbrand Ceder. A first order mott transition in Li_xCoO_2. *Nature Materials*, 3:627–631, 2004.

[25] Ying Shirley Meng, Anton van der Ven, Maria Chan, and Gerbrand Ceder. Ab initio study of sodium ordering in $Na_{0.75}CoO_2$ and its relation to Co3+/Co4+ charge ordering. *Physical Review B*, 72:172103, 2005.

[26] Jorge Nocedal and Stephen Wright. *Numerical Optimization (Springer Series in Operations Research and Financial Engineering).* Springer-Verlag, 2006.

[27] Lars Onsager. Crystal statistics i. a two-dimensional model with a order-disorder transition. *Physical Review*, 65:117–149, 1994.

[28] Michel Roger, Jonathan Morris, Alan Tennant, Jon Goff, Matthias Gutmann Jens-Uwe Hoffmann, Ralf Feyerherm, Dharmalingan Prabhakaran, Nic Shannon, Bella Lake, and Pascale Deen. Patterning of sodium ions and the control of electrons in sodium cobaltate. *Nature*, 445/8:631–634, 2007.

[29] Juan M. Sanchez, Francois Ducastelle, and Denis Gratias. Generalized cluster description of multicomponent systems. *Physica A*, 128:334, 1984.

[30] Raymond Schaak, Tomasz Klimczuk, Maw Lin Foo, and Robert J. Cava. Superconductivity phase diagram of $Na_xCoO_2.1.3H_2O$. *Nature*, 424:527–529, 2003.

[31] Marcel H. F. Sluiter and Yoshiyuki Kawazoe. Invariance of truncated cluster expansions for first-principles alloy thermodynamics. *Physical Review B*, 71:212201, 2005.

[32] Kazunori Takada, Hiroya Sakurai, Eiji Takayama-Muromachi, Fujio Izumi, Ruben A. Dilanian, and Takayoshi Sasaki. Superconductivity in two dimensional CoO_2 layers. *Nature*, VOL 422:53–55, 2003.

[33] Kunio Takezawa. *Introduction to Nonparametric Regression (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 2005.

[34] Ichiro Terasaki, Yoshitaka Sasago, and Kunimitsu Uchinokura. Large thermoelectric power in $NaCo_2O_4$ single crystals. *Physical Review B*, 56:12685–12687, 1997.

[35] Andrei Nikolaevich Tikhonov and Vasilii Arsenin. *Solutions of Ill Posed Problems (Scripta Series in Mathematics)*. Vh Winston Company, 1977.

[36] Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*. SIAM Publications, 1997.

[37] Axel van de Walle and Gerbrand Ceder. Automating first principles phase diagram calculations. *Journal of Phase Equilibria*, 23-4:348–359, 2002.

[38] David H. Wolpert and William G. Macready. No free lunch theorems for optimization. *IEEE Transcations on Evolutionary Computation*, 1:67–82, 1996.

[39] Kasoku Yoshida. *Functional Analysis*. Springer-Verlag, 1980.

[40] Peihong Zhang, Rodrigo B. Capaz, Marvin L. Cohen, and Steven G. Louie. Theory of sodium ordering in Na_xCoO_2. *Physical Review B*, 71:153102, 2005.